THE VARIATION METHOD - II *

by

Saul T. Epstein

Theoretical Chemistry Institute

and

Physics Department

University of Wisconsin

Madison, Wisconsin 53706

(I would be most grateful for any comments, or any suggestions for improvement).

* Research supported by Grant GP-12832 from the National Science Foundation and National Aeronautics and Space Administration Grant NGL 50-002-001.

FACILITY FORM 602

N71-75258 (ACCESSION NUMBER)

137 (PAGES)

CR-122218 (NASA CR OR TMX OR AD NUMBER)

(THRU)

None (CODE)

(CATEGORY)

XXIII. THE VARIATION PRINCIPLE AND PERTURBATION THEORY

There are many connections between the variation method and perturbation theory. In the succeeding sections we will be concerned in the main with using the variation method to approximate solutions of the perturbation equations, and, before that, with the closely related problem of the perturbation analysis of optimal trial functions and energies within the variation method. However the variation principle can play other roles. For example it can often offer insight into exact perturbation theory results. Thus let ν be the perturbation parameter. Then if $\tilde{\psi} - \psi$ is of order ν^{n+1} where ψ is an eigenfunction (not necessarily of the lowest state), it is an immediate consequence of the variation principle that $\tilde{E} - E$ is of order ν^{2n+2} and we have the well known result[1] that to calculate the energy through order ν^{2n+1} one needs to know the wave function only through order ν^n ; a theorem which is usually proven by clever and intricate manipulations of the perturbation equations. In Sec. XXVIII we will present a generalization of this theorem.

In this section we wish to discuss yet another kind of connection; one which , so to speak, goes from perturbation theory to a variation principle rather than the other way round. The point is simply the following. The mark of a variation principle is that it leads to errors of second order in the quantity to be calculated (we have been concentrating on the energy). This suggests that if using perturbation theory we can give a formula for calculating a quantity correct through first order, i.e., so that the error is of second order, then from this formula we should

be able to infer a variation principle for the quantity in question. (This also suggests that starting from the results of higher order perturbation theory one could infer "super-variation principles"[2]).

Rather than discussing in general how to carry out this program[3] we will simply illustrate the point by some examples. First let us consider the energy itself. Then given a zero order wave function $\psi^{(0)}$, a zero order energy $E^{(0)}$, and a zero order Hamiltonian $H^{(0)}$ such that

$$(H^{(0)} - E^{(0)})\,\psi^{(0)} = 0 \qquad \text{(XXIII-1)}$$

one has the familiar result that through first-order the energy eigenvalue for a Hamiltonian H is given by

$$E^{(0)} + \frac{(\psi^{(0)}, (H - H^{(0)})\psi^{(0)})}{(\psi^{(0)}, \psi^{(0)})} \qquad \text{(XXIII-2)}$$

However using (1), the result (2) can be written as

$$(\psi^{(0)}, H\psi^{(0)}) / (\psi^{(0)}, \psi^{(0)}) \qquad \text{(XXIII-3)}$$

In this last form all explicit reference to $H^{(0)}$ has disappeared. Further it is a formula which is correct through first-order in the error in $\psi^{(0)}$, whatever $\psi^{(0)}$ may be. Hence we can infer that

$$(\tilde{\psi}, H\tilde{\psi}) / (\tilde{\psi}, \tilde{\psi})$$

will yield a variational approximation to the energy; and we are back where we started from at the beginning of Section II.

Our second example we will take from an area which we have not discussed heretofore, namely scattering theory. Consider a particle of mass M and energy E which is being scattered by a central field. From the

first Born approximation (i.e., first-order perturbation theory as in (2)) for a single partial wave it then follows that

$$\Gamma^{(0)} + 2M \int_0^\infty dt\, u^{(0)} (h^{(0)} - h) u^{(0)} \qquad \text{(XXIII-4)}$$

yields

$$\Gamma \equiv K^{-1} \tan\eta \qquad \text{(XXIII-5)}$$

correct through first-order when the potential is varied from $V^{(0)}$ to V. Here $K = (2ME/\hbar^2)^{\frac{1}{2}}$, η is the phase shift, $u^{(0)}$ is defined by

$$(h^{(0)} - E) u^{(0)} = 0 \qquad \text{(XXIII-6)}$$

$$u^{(0)} \xrightarrow[t\to\infty]{} \frac{1}{K} \sin(Kt - \ell\tfrac{\pi}{2}) + \Gamma^{(0)} \cos(Kt - \ell\tfrac{\pi}{2}) \qquad \text{(XXIII-7)}$$

and $h^{(0)}$ and h are the radial Hamiltonians for angular momentum ℓ.

We now note that we can eliminate explicit reference to $h^{(0)}$ in (4) by use of (6), and therefore infer that

$$\tilde{\Gamma} + 2M \int_0^\infty dt\, \tilde{u} (E - h) \tilde{u} \qquad \text{(XXIII-8)}$$

will yield a variational approximation to Γ if $\tilde{u}$ is constrained to have the asymptotic form

$$\tilde{u} \xrightarrow[t\to\infty]{} \frac{1}{K} \sin(Kt - \ell\tfrac{\pi}{2}) + \tilde{\Gamma} \cos(Kt - \ell\tfrac{\pi}{2}) \qquad \text{(XXIII-9)}$$

This result is due to W. Kohn[4]. The actual equation of the Kohn variation method is then

$$\delta\hat{\Gamma} + 2M\int_0^\infty dt\ \delta\hat{u}(E-h)\hat{u} + 2M\int_0^\infty dt\ \hat{u}(E-h)\delta\hat{u} = 0 \qquad \text{(XXIII-10)}$$

It is now of some interest to remark[5] that if one does an integration by parts taking due account of the fact that from (9)

$$\delta\hat{u} \xrightarrow[t\to\infty]{} \delta\hat{\Gamma}\cos(kt - \ell\underset{\sim}{\pi})$$

then one finds that (10) can be written

$$\int_0^\infty dt\ \delta\hat{u}(E-h)\hat{u} + \int_0^\infty dt\ [(E-h)\hat{u}]\delta\hat{u} = 0 \qquad \text{(XXIII-11)}$$

which, of course, is completely analogous to equation (V-1).

XXIV. PERTURBATION ANALYSIS OF THE VARIATION METHOD, INTRODUCTION

Let us suppose that our Hamiltonian takes the form

$$H = H^{(0)} + \nu H^{(1)} + \nu^2 H^{(2)} + \cdots \qquad \text{(XXIV-1)}$$

where the $H^{(n)}$ are independent of the real parameter ν, and that we are interested in exhibiting the $\hat{E}$ and other expectation values as power series in ν. Of course such expansions may not exist, but assuming that they do one way to derive them is to first carry out the complete calculation and then expand the results. The other approach, and the one we will explore in this and the next few sections, is to try to produce the successive terms in the series by a direct, iterative method.

The latter has long been the traditional procedure for calculating the response to external fields; however in recent years there has been a trend the other way.[1] Thus to calculate a polarizability, instead of

directly calculating a second order contribution to the energy, one calculates the total energy in a series of weak fields and then fits the results to a polynomial in the field strength. Curiously however, in the calculation of molecular force constants, there has been somewhat of a trend in the other direction,[2] the curve fitting approach having been the more traditional.

Before getting involved in the detailed formalism we should first say a few words about the various possibilities for the $H^{(n)}$ that one encounters in applications. First of all $H^{(0)}$ will usually be either the "accurate" Hamiltonian for an isolated system or some approximation to it. In saying this we qualified the word accurate by using quotation marks because it is a relative term. Thus for an isolated atom or molecule, under many circumstances the fixed nucleus, non-relativistic (no magnetic effects) Hamiltonian is the standard of accuracy, with corrections adequately taken care of by first-order perturbation theory (and the adiabatic approximation for nuclear motion in the case of molecules). However in other circumstances this may well not be the case and one may wish to include relativistic and/or nuclear motion effects in $H^{(0)}$.

If $H^{(0)}$ is an accurate Hamiltonian than the perturbation will usually represent the effects of external field and/or, in molecular applications, the effects of changing the nuclear configuration. On the other hand if $H^{(0)}$ is approximate then the perturbation may also include corrections for the difference between $H^{(0)}$ and the "exact" Hamiltonian of the isolated system. Further the $H^{(n)}$ might include quantities which have been introduced as fictitious perturbations in order to calculate

more accurate expectation values by use of interchange theorems. Thus for complete generality we should really replace each $v^n H^{(n)}$ by a sum of terms. However since one can generalize the formulae at any time by the replacements

$$vH^{(n)} \longrightarrow v_1 H_1^{(n)} + v_2 H_2^{(n)} + \cdots \tag{XXIV-2}$$

$$v^2 H^{(n)} \longrightarrow v_1^2 H_{11}^{(n)} + 2 v_1 v_2 H_{12}^{(n)} + \cdots + v_2^2 H_{22}^{(n)} + 2 v_2 v_3 H_{23}^{(n)} + \cdots \tag{XXIV-3}$$

etc, we will forego this further notational complication for the present (the factors of 2 were introduced into (3) so that $H_{ab}^{(2)}$ is $\left(\partial^2 H / \partial v_a \partial v_b \right)_{v_1 = v_2 = \cdots = 0}$).

In general there are two main (though not always distinct) reasons for wanting to make the expansion of $\hat{E}$ and $\hat{\Psi}$ at all. In the one case, if the perturbations are due to external fields, or to changes in other parameters describing the system then the individual expansion coefficients in $\hat{E}$ will often have a direct physical significance as approximation to polarizabilities, susceptibilities, force constants, etc.[A] In the second case one may be using the expansion as a numerical technique for approximating the solution of the variation equations.[B] A situation in which these two aspects merge is in the use of expansions in $1/\mathcal{Z}_e$ for atoms. On the one hand $1/\mathcal{Z}_e$ is a real physical parameter (albeit not continuously variable) and the coefficients in the series for $\hat{E}$ and other expectation values can be compared with data on isoelectronic series. On the other hand the $1/\mathcal{Z}_e$ expansion also often offers a rapidly converging expansion technique for solving the variation equations for particular values of $\mathcal{Z}_e$.

In what follows our language will almost always be appropriate only to effectively non degenerate situations. However many of the general results are independent of possible complication due to degeneracy and hopefully the reader will be able to tell when this is the case.[3] Thus we will talk as though the zero order equations ((XXV-14) and (XXV-16) below) completely determine the zero order quantities, as though the first-order equations ((XXV-15) and (XXV-17) below) completely determine the first-order quantities etc. Also in our discussion of multiple perturbations we will be quite explicitly dealing with non degenerate situations since we will assume that $\hat{E}$ AND $\hat{\Psi}$ can be expanded in a multiple power series in the ν_α .

XXV. PERTURBATION ANALYSIS OF THE VARIATION METHOD - GENERAL FORMALISM

Let us now return to our problem, the expansion of $\hat{E}$ and $\hat{\Psi}$ in powers of ν . In general $\hat{\Psi}$ will depend on ν , first because the optimal variational parameters and/or functions $\hat{A}$ depend on ν , and in addition there may be some further explicit ν dependence; thus

$$\hat{\Psi} = \Theta(\hat{A}, \nu) \qquad \text{(XXV-1)}$$

We will now assume that the $\hat{A}$ can be expanded in a power series in ν :

$$\hat{A} = \hat{A}^{(0)} + \nu \hat{A}^{(1)} + \nu^2 \hat{A}^{(2)} + \cdots \qquad \text{(XXV-2)}$$

However in Sec. XXVII we will mention a case in which (2) does not hold although still $\hat{\Psi}$ and $\hat{E}$ can be expanded in powers of ν.

Using (2) in (1) we can now expand $\hat{\Psi}$ in a power series in ν as

$$\hat{\Psi} = \hat{\Psi}^{(0)} + \nu \hat{\Psi}^{(1)} + \nu^2 \hat{\Psi}^{(2)} + \cdots \tag{XXV-3}$$

where

$$\hat{\Psi}^{(0)} = \Theta(\hat{A}^{(0)}, 0) \tag{XXV-4}$$

$$\hat{\Psi}^{(1)} = \frac{\partial \hat{\Psi}^{(0)}}{\partial \hat{A}^{(0)}} \hat{A}^{(1)} + \left(\frac{\partial \hat{\Psi}}{\partial \nu}\right)_0 \tag{XXV-5}$$

$$\hat{\Psi}^{(2)} = \frac{\partial \hat{\Psi}^{(0)}}{\partial \hat{A}^{(0)}} \hat{A}^{(2)} + \frac{1}{2} \frac{\partial^2 \hat{\Psi}^{(0)}}{\partial \hat{A}^{(0)2}} (\hat{A}^{(1)})^2 + \left(\frac{\partial^2 \hat{\Psi}}{\partial \nu \partial \hat{A}^{(0)}}\right)_0 \hat{A}^{(1)} + \frac{1}{2}\left(\frac{\partial^2 \hat{\Psi}}{\partial \nu^2}\right)_0 \tag{XXV-6}$$

etc. The subscript 0 means that the quantity is to be evaluated at $\nu = 0$, $\hat{A} = \hat{A}^{(0)}$, and in writing these equations we have used a compact notation, thus

$$\frac{\partial \hat{\Psi}^{(0)}}{\partial \hat{A}^{(0)}} \hat{A}^{(1)} \equiv \sum_i \frac{\partial \hat{\Psi}^{(0)}}{\partial \hat{A}^{(0)}_i} \hat{A}^{(1)}_i \tag{XXV-7}$$

etc.

For $\delta \hat{\Psi}$ we have

$$\delta \hat{\Psi} = \frac{\partial \hat{\Psi}}{\partial \hat{A}} \delta \hat{A} = \frac{\partial \hat{\Psi}}{\partial \hat{A}^{(0)}} \delta \hat{A} \tag{XXV-8}$$

where the $\delta \hat{A}$ are arbitrary within the bounds of whatever the $\hat{A}$ are - numerical parameters and/or arbitrary functions of certain variables.

Then using equations (3) through (6) we find

$$\delta\hat{\psi} = \delta^0\hat{\psi} + \nu\delta^1\hat{\psi} + \nu^2\delta^2\hat{\psi} + \cdots \qquad \text{(XXV-9)}$$

where

$$\delta^0\hat{\psi} = \frac{\partial\hat{\psi}^{(0)}}{\partial\hat{A}^{(0)}}\delta\hat{A} \qquad \text{(XXV-10)}$$

$$\delta^1\hat{\psi} = \frac{\partial^2\hat{\psi}^{(0)}}{\partial\hat{A}^{(0)2}}\delta\hat{A}\,\hat{A}^{(1)} + \left(\frac{\partial^2\hat{\psi}}{\partial\hat{A}^{(0)}\partial\nu}\right)_0\delta\hat{A} \qquad \text{(XXV-11)}$$

etc.

We now insert these results into the basic equations of the variation method

$$(\delta\hat{\psi}, (H-\hat{E})\hat{\psi}) + ((H-\hat{E})\hat{\psi}, \delta\hat{\psi}) = 0 \quad \text{"all" } \delta\hat{\psi} \qquad \text{(XXV-12)}$$

and

$$(\hat{\psi}, (H-\hat{E})\hat{\psi}) = 0 \qquad \text{(XXV-13)}$$

Then writing $\hat{E}$ as

$$\hat{E} = \hat{E}^{(0)} + \nu\hat{E}^{(1)} + \nu^2\hat{E}^{(2)} + \cdots$$

and equating separately to zero the terms of each order in ν, (12) yields

$$(\delta^0\hat{\psi}, (H^{(0)} - \hat{E}^{(0)})\hat{\psi}^{(0)}) + \text{c.c.} = 0 \quad \text{All } \delta\hat{A} \qquad \text{(XXV-14)}$$

$$\{(\delta^0\hat{\psi}, (H^{(0)} - \hat{E}^{(0)})\hat{\psi}^{(1)}) + (\delta^0\hat{\psi}, (H^{(1)} - \hat{E}^{(1)})\hat{\psi}^{(0)}) +$$

$$+ (\delta^1\hat{\psi}, (H^{(0)} - \hat{E}^{(0)})\hat{\psi}^{(0)})\} + \text{c.c.} = 0 \qquad \text{(XXV-15)}$$

etc.

Similarly (13) yields

$$(\hat{\Psi}^{(0)}, (H^{(0)} - \hat{E}^{(0)})\hat{\Psi}^{(0)}) = 0 \qquad \text{(XXV-16)}$$

$$(\hat{\Psi}^{(1)}, (H^{(0)} - \hat{E}^{(0)})\hat{\Psi}^{(0)}) + (\hat{\Psi}^{(0)}, (H^{(0)} - \hat{E}^{(0)})\hat{\Psi}^{(1)}) + (\hat{\Psi}^{(0)}, (H^{(1)} - \hat{E}^{(1)})\hat{\Psi}^{(0)}) = 0 \qquad \text{(XXV-17)}$$

$$(\hat{\Psi}^{(2)}, (H^{(0)} - \hat{E}^{(0)})\hat{\Psi}^{(0)}) + (\hat{\Psi}^{(0)}, (H^{(0)} - \hat{E}^{(0)})\hat{\Psi}^{(2)}) + (\hat{\Psi}^{(1)}, (H^{(0)} - \hat{E}^{(0)})\hat{\Psi}^{(1)}) +$$

$$+ (\hat{\Psi}^{(1)}, (H^{(1)} - \hat{E}^{(1)})\hat{\Psi}^{(0)}) + (\hat{\Psi}^{(0)}, (H^{(1)} - \hat{E}^{(1)})\hat{\Psi}^{(1)}) +$$

$$+ (\hat{\Psi}^{(0)}, (H^{(2)} - \hat{E}^{(2)})\hat{\Psi}^{(0)}) = 0$$

(XXV-18)

We will discuss these equations in detail only through first-order in the wave function and second-order in the energy since this will suffice to illustrate most of the essential features.

Equations (14) and (16) need little comment. They are just the basic equations of the variation method for determining $\hat{\Psi}^{(0)}$ and $\hat{E}^{(0)}$. Turning now to (15), if we insert $\delta''\hat{\Psi}$ and $\delta'\hat{\Psi}$ from (10) and (11) and $\hat{\Psi}^{(1)}$ from (5) then we see that (15) will yield equations to determine the $\hat{A}^{(1)}$. In particular if the $\hat{A}$ are simple numerical parameters (linear or not) then (15) will yield a set of linear inhomogeneous equations to determine the $\hat{A}^{(1)}$. Also, clearly, a similar result will hold in higher orders.[1]

Having determined the $\hat{A}^{(0)}$ and $\hat{A}^{(1)}$ equation (17) allows us to determine $\hat{E}^{(1)}$. Indeed we can now also determine $\hat{E}^{(2)}$ from

(18), although at first sight this seems to require a knowledge of $\hat{A}^{(2)}$. However, from (18) and (6), the $\hat{A}^{(2)}$ contribution appears in the form

$$\left(\frac{\partial \hat{\psi}^{(0)}}{\partial \hat{A}^{(0)}} \hat{A}^{(2)}, \; (H^{(0)} - \hat{E}^{(0)}) \hat{\psi}^{(0)} \right) + c.c. \qquad \text{(XXV-19)}$$

which in turn is of the form

$$\frac{1}{\delta \epsilon} \left(\delta^0 \hat{\psi}, (H^{(0)} - \hat{E}^{(0)}) \hat{\psi}^{(0)} \right) + c.c.$$

with

$$\delta \hat{A} = \delta \epsilon \hat{A}^{(2)}$$

where $\delta\epsilon$ is an infinitesimal real number. Therefore from (14) it follows that the contribution (19) vanishes.

Thus we can replace (18) by

$$\begin{aligned} &(\hat{\psi}^{(2)\prime}, (H^{(0)} - \hat{E}^{(0)}) \hat{\psi}^{(0)}) + (\hat{\psi}^{(0)}, (H^{(0)} - \hat{E}^{(0)}) \hat{\psi}^{(2)\prime}) + (\hat{\psi}^{(1)}, (H^{(0)} - \hat{E}^{(0)}) \hat{\psi}^{(1)\prime}) \\ &+ (\hat{\psi}^{(1)}, (H^{(1)} - \hat{E}^{(1)}) \hat{\psi}^{(0)}) + (\hat{\psi}^{(0)}, (H^{(1)} - \hat{E}^{(1)}) \hat{\psi}^{(1)}) \\ &+ (\hat{\psi}^{(0)}, (H^{(0)} - \hat{E}^{(0)}) \hat{\psi}^{(0)}) = 0 \end{aligned} \qquad \text{(XXV-20)}$$

where

$$\hat{\psi}^{(2)\prime} = \frac{1}{2} \frac{\partial^2 \hat{\psi}^{(0)}}{\partial \hat{A}^{(0)\nu} \partial \hat{A}^{(0)\nu}} (\hat{A}^{(1)})^2 + \left(\frac{\partial^2 \hat{\psi}}{\partial \nu \partial \hat{A}^{(0)}} \right)_0 \hat{A}^{(1)} + \frac{1}{2} \left(\frac{\partial^2 \hat{\psi}}{\partial \nu^2} \right)_0 \qquad \text{(XXV-21)}$$

and we have the result that if we know $\hat{A}^{(0)}$ and $\hat{A}^{(1)}$ then we can calculate $\hat{E}^{(2)}$. In Sec. XXVII we will shown, as part of a more general discussion, that this information is also sufficient to calculate $\hat{E}^{(3)}$.

Let us now examine the formula (17) for $\hat{E}^{(1)}$ more closely. First of all we note that actually the $\hat{A}^{(1)}$ are not needed to calculate $\hat{E}^{(1)}$ since, from (5) and (17) the contribution of the $\hat{A}^{(1)}$ terms is in the form

$$\left(\frac{\partial \hat{\Psi}^{(0)}}{\partial \hat{A}^{(0)}} \hat{A}^{(1)}, (H^{(0)} - \hat{E}^{(0)}) \hat{\Psi}^{(0)} \right) + c.c.$$

which, from the previous argument, vanishes. Thus we have the general result

$$\hat{E}^{(1)} = \frac{(\hat{\Psi}^{(0)}, H^{(1)} \hat{\Psi}^{(0)})}{(\hat{\Psi}^{(0)}, \hat{\Psi}^{(0)})} + \frac{\left(\left(\frac{\partial \hat{\Psi}}{\partial \nu} \right)_0, (H^{(0)} - \hat{E}^{(0)}) \hat{\Psi}^{(0)} \right) + c.c.}{(\hat{\Psi}^{(0)}, \hat{\Psi}^{(0)})} \qquad \text{(XXV-22)}$$

In Sec. XVI we showed that if the set of trial functions is independent of ν then the generalized Hellmann-Feynman theorem for $\sigma = \nu$ will be satisfied. Thus

$$\frac{\partial \hat{E}}{\partial \nu} = \frac{(\hat{\Psi}, (H^{(1)} + 2\nu H^{(2)} + \cdots) \hat{\Psi})}{(\hat{\Psi}, \hat{\Psi})} \qquad \text{(XXV-23)}$$

and hence, since

$$\hat{E}^{(1)} = \left(\frac{\partial \hat{E}}{\partial \nu} \right)_0$$

we have, instead of (22), the simpler formula

$$\hat{E}^{(1)} = \frac{(\hat{\Psi}^{(0)}, H^{(1)} \hat{\Psi}^{(0)})}{(\hat{\Psi}^{(0)}, \hat{\Psi}^{(0)})} \qquad \text{(XXV-24)}$$

To gain further insights into the workings of perturbation theory within the variation method we will now rederive (24) directly from the detailed equations of the present section by imposing the further require-

ment that the set of $\tilde{\Psi}$ be independent of ν. In some cases like UHF, OHF, and the linear variation method with the individual basis function independent of ν, the set is invariant to changes in ν because the individual members are independent of ν. Thus in such a case (1) becomes $\hat{\Psi} = \Theta(\hat{A})$ and we have

$$\frac{\partial \hat{\Psi}}{\partial \nu} \equiv 0 \tag{XXV-25}$$

so that (24) then follows trivially from (22). However more generally, invariance of the set to changes in ν implies that a change in ν is equivalent to a change in the $\tilde{A}$. Thus we have

$$\Theta(\hat{A}, \nu + \Delta\nu) = \Theta(\hat{A} + \Delta, \nu) \tag{XXV-26}$$

where the exact nature of Δ will depend on the nature of the trial functions. Assuming however that Δ can be written

$$\Delta = C\,\Delta\nu + D(\Delta\nu)^2 + \cdots \tag{XXV-27}$$

where C, D, ... are independent of $\Delta\nu$ we find, by expanding (26) in powers of $\Delta\nu$, that

$$\frac{\partial \hat{\Psi}}{\partial \nu} = \frac{\partial \hat{\Psi}}{\partial \hat{A}} C \tag{XXV-28}$$

$$\frac{\partial^2 \hat{\Psi}}{\partial \nu^2} = \frac{\partial^2 \hat{\Psi}}{\partial \hat{A}^2} C^2 + 2D \frac{\partial \hat{\Psi}}{\partial \hat{A}} \tag{XXV-29}$$

etc. Further differentiating (28) with respect to $\hat{A}$ we have

$$\frac{\partial^2 \hat{\Psi}}{\partial \hat{A}\,\partial \nu} = \frac{\partial^2 \hat{\Psi}}{\partial \hat{A}^2} C + \frac{\partial \hat{\Psi}}{\partial \hat{A}} \frac{\partial C}{\partial \hat{A}} \tag{XXV-30}$$

To derive (24) from (22) we need note only that (28) yields

$$\left(\frac{\partial\hat{\psi}}{\partial\nu}\right)_0 = \frac{\partial\hat{\psi}^{(0)}}{\partial\hat{A}^{(0)}}\, C_0 \qquad \text{(XXV-31)}$$

so that, by the argument which we have used twice before,

$$\left(\left(\frac{\partial\hat{\psi}}{\partial\nu}\right)_0 , (H - \hat{E}^{(0)})\hat{\psi}^{(0)} \right) + \text{c.c.} = 0 \qquad \text{(XXV-32)}$$

Returning to the formula (20) for $\hat{E}^{(1)}$, and again with the excuse that we will thereby gain further insight and practice in the use of the machinery, let us specialize it to the case in which (26) is satisfied. Writing (23) in the form

$$\left(\hat{\psi}, \left(\frac{\partial\hat{E}}{\partial\nu} - H^{(0)} - 2\nu H^{(1)} - \cdots\right)\hat{\psi}\right) = 0 \qquad \text{(XXV-33)}$$

and differentiating again, we see that we should find

$$\hat{E}^{(1)} \equiv \frac{1}{2}\left(\frac{\partial^2\hat{E}}{\partial\nu^2}\right)_0 = \frac{(\hat{\psi}^{(0)}, H^{(1)}\hat{\psi}^{(0)})}{(\hat{\psi}^{(0)}, \hat{\psi}^{(0)})} + \frac{1}{2}\frac{(\hat{\psi}^{(0)}, (H^{(2)} - \hat{E}^{(2)})\hat{\psi}^{(0)}) + \text{c.c.}}{(\hat{\psi}^{(0)}, \hat{\psi}^{(0)})} \qquad \text{(XXV-34)}$$

From (29) and (30) we have

$$\left(\frac{\partial^2\hat{\psi}}{\partial\nu^2}\right)_0 = \frac{\partial^2\hat{\psi}^{(0)}}{\partial\hat{A}^{(0)2}}\, C_0^2 + 2D_0\frac{\partial\hat{\psi}^{(0)}}{\partial\hat{A}^{(0)}} \qquad \text{(XXV-35)}$$

and

$$\left(\frac{\partial^2\hat{\psi}}{\partial\hat{A}\,\partial\nu}\right)_0 = \frac{\partial^2\hat{\psi}^{(0)}}{\partial\hat{A}^{(0)2}}\, C_0 + \frac{\partial\hat{\psi}^{(0)}}{\partial\hat{A}^{(0)}}\left(\frac{\partial C}{\partial\hat{A}}\right)_0 \qquad \text{(XXV-36)}$$

If now we use (31), (35), and (36) in (5), (21), and (11) then we find

$$\hat{\psi}^{(1)} = \frac{\partial\hat{\psi}^{(0)}}{\partial\hat{A}^{(0)}}\, G \qquad \text{(XXV-37)}$$

$$\hat{\Psi}^{(2)'} = \frac{1}{2}\frac{\partial^2\hat{\Psi}^{(0)}}{\partial\hat{A}^{(0)2}}(G)^2 + \frac{\partial\hat{\Psi}^{(0)}}{\partial\hat{A}^{(0)}}\left[D_0 + \left(\frac{\partial C}{\partial\hat{A}}\right)_0\hat{A}^{(1)}\right] \qquad \text{(XXV-38)}$$

and

$$\delta^1\hat{\Psi} = \frac{\partial^2\hat{\Psi}^{(0)}}{\partial\hat{A}^{(0)2}}G\,\delta\hat{A} + \frac{\partial\hat{\Psi}^{(0)}}{\partial\hat{A}^{(0)}}\left(\frac{\partial C}{\partial\hat{A}}\right)_0\delta\hat{A} \qquad \text{(XXV-39)}$$

where

$$G \equiv \hat{A}^{(1)} + C_0$$

Therefore, by our standard argument, we can drop the last term on the right in (38) in calculating $\hat{E}^{(2)}$ so that we have from (20)

$$\left[\left(\frac{1}{2}\frac{\partial^2\hat{\Psi}^{(0)}}{\partial\hat{A}^{(0)2}}G^2,\,(H^{(0)}-\hat{E}^{(0)})\hat{\Psi}^{(1)}\right)+c.c.\right] + \left(\frac{\partial\hat{\Psi}^{(0)}}{\partial\hat{A}^{(0)}}G,\,(H^{(0)}-\hat{E}^{(0)})\frac{\partial\hat{\Psi}^{(0)}}{\partial\hat{A}^{(0)}}G\right)$$

$$+\left[\left(\frac{\partial\hat{\Psi}^{(0)}}{\partial\hat{A}^{(0)}}G,\,(H^{(1)}-\hat{E}^{(1)})\hat{\Psi}^{(0)}\right)+c.c.\right] + \left(\hat{\Psi}^{(0)},\,(H^{(2)}-\hat{E}^{(2)})\hat{\Psi}^{(0)}\right)=0 \qquad \text{(XXV-40)}$$

On the other hand if in (15) we choose

$$\delta\hat{A} = \delta\epsilon\, G$$

and note that we can drop the last term on the right in (39) in evaluating (15) then we find that

$$2\left(\frac{\partial\hat{\Psi}^{(0)}}{\partial\hat{A}^{(0)}}G,\,(H^{(0)}-\hat{E}^{(0)})\frac{\partial\hat{\Psi}^{(0)}}{\partial\hat{A}^{(0)}}G\right) + \left[\left(\frac{\partial\hat{\Psi}^{(0)}}{\partial\hat{A}^{(0)}}G,\,(H^{(1)}-\hat{E}^{(1)})\hat{\Psi}^{(0)}\right)+c.c.\right]$$

$$+\left[\left(\frac{\partial^2\hat{\Psi}^{(0)}}{\partial\hat{A}^{(0)2}}G^2,\,(H^{(0)}-\hat{E}^{(0)})\hat{\Psi}^{(0)}\right)+c.c.\right]=0 \qquad \text{(XXV-41)}$$

Combining (40) and (41) then yields

$$\frac{1}{2}\left[\left(\frac{\partial\hat{\Psi}^{(0)}}{\partial\hat{A}^{(1)}}\,G,\,(H^{(2)}-\hat{E}^{(1)})\hat{\Psi}^{(1)}\right)+\text{c.c.}\right]+(\hat{\Psi}^{(0)},(H^{(2)}-\hat{E}^{(2)})\hat{\Psi}^{(1)})=0 \qquad \text{(XXV-42)}$$

which, using (37) again, is (34).

Variational approximations in which the generalized Hellmann-Feynman theorem is satisfied for $\sigma=\nu$ also have other special and useful properties. For example we will now show[2] that if $\hat{E}^{(0)}$ is a local minimum then

$$\hat{E}^{(2)} \le \frac{(\hat{\Psi}^{(0)},\,H^{(2)}\hat{\Psi}^{(0)})}{(\hat{\Psi}^{(0)},\,\hat{\Psi}^{(0)})} \qquad \text{(XXV-43)}$$

To prove this we use (41) to rewrite (42) in the form

$$\hat{E}^{(2)}-\frac{(\hat{\Psi}^{(0)},H^{(2)}\hat{\Psi}^{(0)})}{(\hat{\Psi}^{(0)},\hat{\Psi}^{(0)})}=-\frac{1}{2(\hat{\Psi}^{(0)},\hat{\Psi}^{(0)})}\left[\left(\frac{\partial^2\hat{\Psi}^{(0)}}{\partial\hat{A}^{(1)\nu}}G^2,(H^{(0)}-\hat{E}^{(0)})\hat{\Psi}^{(1)}\right)+\right.$$

$$\left.+2\left(\frac{\partial\hat{\Psi}^{(0)}}{\partial\hat{A}^{(1)\nu}}G,(H^{(0)}-\hat{E}^{(0)})\frac{\partial\hat{\Psi}^{(0)}}{\partial\hat{A}^{(1)\nu}}G\right)+\left(\hat{\Psi}^{(0)},(H^{(0)}-\hat{E}^{(0)})\frac{\partial^2\hat{\Psi}^{(0)}}{\partial\hat{A}^{(1)\nu}}G^2\right)\right] \qquad \text{(XXV-44)}$$

However starting from

$$(\hat{\Psi}^{(0)},\,(H^{(0)}-\hat{E}^{(0)})\,\hat{\Psi}^{(0)}) \qquad \text{(XXV-45)}$$

we find, since $\delta\hat{E}^{(0)}=0$, that in general

$$\delta^2\hat{E}^{(0)}(\hat{\Psi}^{(0)},\hat{\Psi}^{(0)})=\left[\left(\frac{\partial^2\hat{\Psi}^{(0)}}{\partial\hat{A}^{(1)\nu}}(\delta\hat{A})^\nu,(H^{(0)}-\hat{E}^{(0)})\hat{\Psi}^{(0)}\right)+\right.$$

$$\left.+2\left(\frac{\partial\hat{\Psi}^{(0)}}{\partial\hat{A}^{(1)\nu}}\delta\hat{A},(H^{(0)}-\hat{E}^{(0)})\frac{\partial\hat{\Psi}^{(0)}}{\partial\hat{A}^{(1)\nu}}\delta\hat{A}\right)+\left(\hat{\Psi}^{(0)},(H^{(0)}-\hat{E}^{(0)})\frac{\partial^2\hat{\Psi}^{(0)}}{\partial\hat{A}^{(1)\nu}}(\delta\hat{A})^\nu\right)\right] \qquad \text{(XXV-46)}$$

Therefore comparing the right hand side of (44) with the right hand side of (46), with $\delta\hat{A}$ taken equal to $\delta\epsilon\,G$, we see, as announced, that

if $\delta^2 \hat{E}^{(0)} \neq 0$ (at least for this $\delta \hat{A}$) then we will have (34).

Let us now suppose that there are two perturbations acting (the generalization to more than two being trivial) and let us suppose that $H_1^{(1)}$ etc. are easier perturbations to deal with than $H_2^{(1)}$, etc. Then, assuming no problems with degeneracy so that the double power series exist we will have, in the notation of Eq. (XXIV-2) and (XXIV-3)

$$\hat{E} = \hat{E}^{(0)} + \nu_1 \hat{E}_1^{(1)} + \nu_2 \hat{E}_2^{(1)} + \nu_1^2 \hat{E}_{11}^{(2)} + 2\nu_1\nu_2 \hat{E}_{12}^{(2)} + \nu_2^2 \hat{E}_{22}^{(2)} + \nu_1^3 \hat{E}_{111}^{(3)} + \tfrac{3}{2}\nu_1^2\nu_2 \hat{E}_{112}^{(3)} + \cdots \qquad \text{(XXV-47)}$$

In Section XXVII we will show that in general there are interchange theorems which can aid in the calculation of $\hat{E}$ through order n_1 in ν_1 and order n_2 in ν_2 by minimizing the accuracy needed with respect to ν_2 in the calculation of $\hat{A}$. If the generalized Hellmann-Feynman theorem for $\sigma = \nu_1$ is satisfied

$$\frac{\partial \hat{E}}{\partial \nu_1} = (\psi, \frac{\partial H}{\partial \nu_1} \psi)/(\psi, \psi) \qquad \text{(XXV-48)}$$

then this will immediately enable us also to calculate the average of $\partial H/\partial \nu_1$ correctly through order $(n_1 - 1)$ in ν_1 and n_2 in ν_2 . Also, as we will see, the formulae take a particularly simple form if the generalized Hellmann-Feynman theorem for $\sigma = \nu_2$

$$\frac{\partial \hat{E}}{\partial \nu_2} = (\psi, \frac{\partial H}{\partial \nu_2} \psi)/(\psi, \psi) \qquad \text{(XXV-49)}$$

is also satisfied.

As we said, we will discuss the general situation in Sec. XXVII; here we will discuss only the special case $n_1 = r$, $n_2 = 1$. Also we will assume that (48) and (49) are both satisfied,[3] and, as a further simplification we will assume that all "cross perturbations" vanish, i.e.,

$$H^{(2)}_{12} = H^{(3)}_{112} = \cdots = 0$$

We now want to calculate some, or all, of the quantities

$$\hat{E}^{(0)}, \hat{E}^{(1)}_1, \hat{E}^{(1)}_2, \hat{E}^{(2)}_{11}, \hat{E}^{(2)}_{12}, \hat{E}^{(3)}_{112} \qquad \text{(XXV-50)}$$

From the earlier results in this section we know, identifying ν with ν_1 or ν_2 as is appropriate, that the first three quantities can be calculated if we know $\hat{A}$ through zero order in ν_1 and zero order in ν_2 while the fourth requires $\hat{A}$ through first-order in ν_1 but still zero order in ν_2. We now consider the quantities $\hat{E}^{(2)}_{12}$ and $\hat{E}^{(3)}_{112}$.

In general

$$\hat{E}^{(2)}_{12} = \frac{1}{2}\left(\frac{\partial^2 \hat{E}}{\partial \nu_2 \partial \nu_1}\right)_{\nu_1=0,\nu_2=0} \qquad \text{(XXV-51)}$$

and hence, from (47) and (48), $2\nu_2 \hat{E}^{(2)}_{12}$ yields the first-order in ν_2 correction to the approximation to the expectation value of $H^{(2)}_1$ for $\nu_1 = 0$, furnished by $\hat{\psi}^{(0)}$.

One approach to calculating $\hat{E}^{(2)}_{12}$ is to use (48) and replace (51) by

$$\hat{E}^{(2)}_{12} = \frac{1}{2}\left[\frac{\partial}{\partial \nu_2}\left[\frac{(\psi, H^{(2)}_1 \psi)}{(\psi, \psi)}\right]_{\nu_1=0}\right]_{\nu_2=0} \qquad \text{(XXV-52)}$$

If then we write $\hat{\Psi}$ for $\nu_1=0$ as

$$\hat{\Psi}(\nu_1=0) = \hat{\Psi}^{(0)} + \nu_2 \hat{\Psi}_2^{(1)} + \nu_2^2 \hat{\Psi}_2^{(2)} + \cdots \qquad \text{(XXV-53)}$$

(52) easily yields

$$\hat{E}_{12}^{(2)} = \frac{1}{2} \frac{(\hat{\Psi}^{(0)}, (H_1^{(1)} - \langle H_1^{(1)} \rangle_0)\hat{\Psi}_2^{(1)}) + c.c.}{(\hat{\Psi}^{(0)}, \hat{\Psi}^{(0)})} \qquad \text{(XXV-54)}$$

where

$$\langle H_1^{(1)} \rangle_0 = (\hat{\Psi}^{(0)}, H_1^{(1)} \hat{\Psi}^{(0)}) / (\hat{\Psi}^{(0)}, \hat{\Psi}^{(0)}) = \hat{E}_1^{(1)}$$

Thus in this way of calculating $\hat{E}_{12}^{(2)}$ we must know $\hat{A}$ through only zero order in ν_1 but through first-order in ν_2. Note also that whether or not (48) is satisfied and whether or not the double power series exists we certainly have, from (53), that

$$\left[\frac{(\hat{\Psi}, H_1^{(1)} \hat{\Psi})}{(\hat{\Psi}, \hat{\Psi})} \right]_{\nu_1=0} = \langle H_1^{(1)} \rangle_0 + \nu_2 \frac{(\hat{\Psi}^{(0)}, (H_2^{(1)} - \langle H_2^{(1)} \rangle_0)\hat{\Psi}_2^{(1)}) + c.c.}{(\hat{\Psi}^{(0)}, \hat{\Psi}^{(0)})} + \cdots \qquad \text{(XXV-55)}$$

for any $H_1^{(1)}$.

However we now note that we can, by interchanging the orders of differentiation and using (49), write (52) as

$$\hat{E}_{12}^{(2)} = \frac{1}{2} \left[\frac{\partial}{\partial \nu_1} \left[\frac{(\hat{\Psi}, H_2^{(1)} \hat{\Psi})}{(\hat{\Psi}, \hat{\Psi})} \right]_{\nu_2=0} \right]_{\nu_1=0} \qquad \text{(XXV-56)}$$

which, if we write $\hat{\Psi}$ for $\nu_2=0$ as

$$\hat{\Psi}(\nu_2=0) = \hat{\Psi}^{(0)} + \nu_1 \hat{\Psi}_1^{(1)} + \nu_1^2 \hat{\Psi}_1^{(2)} + \cdots \qquad \text{(XXV-53*)}$$

yields

$$\hat{E}_{12}^{(2)} = \frac{1}{2} \frac{(\hat{\Psi}^{(0)}, (H_2^{(1)} - \langle H_2^{(1)} \rangle_0)\hat{\Psi}_1^{(1)}) + c.c.}{(\hat{\Psi}^{(0)}, \hat{\Psi}^{(0)})} \qquad \text{(XXV-57)}$$

Thus by this interchange theorem we can calculate $\hat{E}^{(2)}_{12}$ by knowing $\hat{A}$ through first-order in ν_1 but only zero order in ν_2. We will now show that we can also calculate $\hat{E}^{(3)}_{112}$ if in addition we know $\hat{A}$ through second-order in ν_1 but still zero order in ν_2. More generally we will show in Section XXVII that whether or not either of (48) or (49) hold, still one can calculate $\hat{E}$ correctly through order M_1 in ν_1 and first-order in ν_2 if one knows $\hat{A}$ correctly through order M_1 in ν_1 and zero order in ν_2.

Returning to $\hat{E}^{(3)}_{112}$ we can write it as

$$\hat{E}^{(3)}_{112} = \frac{1}{3}\left(\frac{\partial}{\partial\nu_2}\left(\frac{\partial^2\hat{E}}{\partial\nu_1^2}\right)_{\nu_1=0}\right)_{\nu_2=0} \qquad \text{(XXV-58)}$$

and hence from (47) $3\nu_2\hat{E}^{(3)}_{112}$ yields a first-order in ν_2 correction to the approximation to the "polarizability" $(\partial^2\hat{E}/\partial\nu_1^2)_{\nu_1=0}$ calculated by assuming $\nu_2=0$. We now note that by interchanging the order of differentiation and using (49) we can rewrite (58) as

$$\hat{E}^{(3)}_{112} = \frac{1}{3}\left(\frac{\partial^2}{\partial\nu_1^2}\left(\frac{(\hat{\Psi}, H^{(1)}_2\hat{\Psi})}{(\hat{\Psi},\hat{\Psi})}\right)_{\nu_2=0}\right)_{\nu_1=0} \qquad \text{(XXV-59)}$$

Then by use of (53*) one finds

$$\hat{E}^{(3)}_{112} = \frac{1}{3}\frac{(\hat{\Psi}^{(0)}_1,(H^{(1)}_2-\langle H^{(1)}_2\rangle_0)\hat{\Psi}^{(1)}_1)+\{(\hat{\Psi}^{(2)}_1,(H^{(1)}_2-\langle H^{(1)}_2\rangle_0)\hat{\Psi}^{(0)})+\text{c.c.}\}}{(\hat{\Psi}^{(0)},\hat{\Psi}^{(0)})}. \qquad \text{(XXV-60)}$$

which proves the point.

If degeneracy is a problem so that the double power series expansion of the energy does not exist (i.e., the limits $\nu_1\to 0$, $\nu_2\to 0$ are not interchangeable) then interchange theories may not exist. Thus for

"first-order properties", i.e., average values evaluated at $\nu_1 = 0$ one, as we said earlier, always has (55) but it may or may not be possible to replace the coefficient of $2\nu_2$ by something like the right hand side of (57).[4]

For "second-order properties", i.e., for $\left(\partial^2 \hat{E}/\partial \nu_1^2\right)_{\nu_1=0}$ the situation is even more complicated in that this may involve inverse powers of ν_ν so that (58) is irrelevant (and infinite). However still it may be possible to use interchange theorems to simplify the calculation of first-order corrections to the leading term,[A]

XXVI. VARIATION METHODS WITHIN THE VARIATION METHOD

Let us denote by $\delta\hat{\Psi}^{(\omega)}$ and $\delta\hat{\Psi}^{(\omega')}$ the results of an arbitrary variation $\delta\hat{A}$ of $\hat{A}^{(\omega)}$ in $\hat{\Psi}^{(\omega)}$ and $\hat{\Psi}^{(\omega')}$ respectively. Then from (XXV-5) and (XXV-20) we have

$$\delta\hat{\Psi}^{(\omega)} = \frac{\partial\hat{\Psi}^{(\omega)}}{\partial\hat{A}^{(\omega)}}\,\delta\hat{A} \qquad \text{(XXVI-1)}$$

$$\delta\hat{\Psi}^{(\omega')} = \frac{\partial^2\hat{\Psi}^{(\omega)}}{\partial\hat{A}^{(\omega)2}}\,\hat{A}^{(\omega)}\,\delta\hat{A} + \left(\frac{\partial^2\hat{\Psi}}{\partial\nu\,\partial\hat{A}^{(\omega)}}\right)_D \delta\hat{A} \qquad \text{(XXVI-2)}$$

Comparing (1) and (2) with (XXV-10) and (XXV-11) then shows that

$$\delta\hat{\Psi}^{(\omega)} = \delta^0\Psi \qquad \text{(XXVI-3)}$$

$$\delta\hat{\Psi}^{(\omega')} = \delta^1\Psi \qquad \text{(XXVI-4)}$$

Therefore with this understanding as to the meaning of δ, eqn. (XXV-15) can be written

$$\delta \hat{J}^{(2)} = 0 \qquad \text{all } \delta \hat{A} \qquad \text{(XXVI-5)}$$

where

$$\hat{J}^{(2)} = (\hat{\psi}^{(1)}, (H^{(0)} - \hat{E}^{(0)})\hat{\psi}^{(1)}) + (\hat{\psi}^{(0)}, (H^{(0)} - \hat{E}^{(0)})\hat{\psi}^{(1)\prime}) + (\hat{\psi}^{(2)\prime}, (H^{(0)} - \hat{E}^{(0)})\hat{\psi}^{(0)}) + (\hat{\psi}^{(0)}, (H^{(1)} - \hat{E}^{(1)})\hat{\psi}^{(1)}) + (\hat{\psi}^{(1)}, (H^{(1)} - \hat{E}^{(1)})\hat{\psi}^{(0)}) \qquad \text{(XXVI-6)}$$

From (XXV-20) we see that the numerical value of $\hat{J}^{(2)}$ is

$$\hat{J}^{(2)} = (\hat{\psi}^{(0)}, (\hat{E}^{(2)} - H^{(2)})\hat{\psi}^{(0)}) \equiv \hat{E}^{(2)\prime}(\hat{\psi}^{(0)}, \hat{\psi}^{(0)}) \qquad \text{(XXVI-7)}$$

and hence yields the "non-trivial" part of $\hat{E}^{(2)}$ since $(\hat{\psi}^{(0)}, H^{(2)}\hat{\psi}^{(0)})$ is readily calculable. Further, if we wish, we can clearly replace $\hat{\psi}^{(2)\prime}$ by $\hat{\psi}^{(2)}$ and can also extend the definition of δ to include variations of $\hat{A}^{(2)}$ since, by the familiar argument the extra contributions to $\hat{J}^{(2)}$ and to $\delta\hat{J}^{(2)}$ will vanish identically.

Now to the immediate purpose of this exercise. As we have seen, Equation (XXV-15) yields equations to determine the $\hat{A}^{(1)}$. However in many practical situations one cannot actually solve these equations. The present considerations then suggest a variational technique for approximating their solution. Namely calculate $\tilde{J}^{(2)}$ defined by

$$\hat{J}^{(2)} = (\hat{\Psi}^{(1)}, (H^{(0)} - \hat{E}^{(0)})\hat{\Psi}^{(1)}) +$$

$$+ (\hat{\Psi}^{(0)}, (H^{(0)} - \hat{E}^{(0)})\hat{\Psi}^{(2)\prime}) + (\hat{\Psi}^{(2)\prime}, (H^{(0)} - \hat{E}^{(0)})\hat{\Psi}^{(0)}) +$$

$$+ (\hat{\Psi}^{(0)}, (H^{(1)} - \hat{E}^{(1)})\hat{\Psi}^{(1)}) + (\hat{\Psi}^{(1)}, (H^{(1)} - \hat{E}^{(1)})\hat{\Psi}^{(0)}) \qquad \text{(XXVI-8)}$$

where

$$\hat{\Psi}^{(1)} = \frac{\partial \hat{\Psi}^{(0)}}{\partial \hat{A}^{(0)}} \tilde{A}^{(1)} + \left(\frac{\partial \Psi}{\partial \nu}\right)_0 \qquad \text{(XXVI-9)}$$

$$\hat{\Psi}^{(2)\prime} = \frac{1}{2} \frac{\partial^2 \hat{\Psi}^{(0)}}{\partial \hat{A}^{(0)2}} (\tilde{A}^{(1)})^2 + \left(\frac{\partial^2 \Psi}{\partial \nu \partial \hat{A}^{(0)}}\right)_0 \tilde{A}^{(1)} + \frac{1}{2} \left(\frac{\partial^2 \Psi}{\partial \nu^2}\right)_0 \qquad \text{(XXVI-10)}$$

and where the $\tilde{A}^{(1)}$ are to be drawn from some limited set, no larger than one can handle. Then require

$$\delta \tilde{J}^{\nu(2)} = 0 \qquad \text{"all" } \delta \tilde{A}^{\nu(1)} \qquad \text{(XXVI-11)}$$

where we use the symbol $^{\nu}$ to denote the optimal quantities which will emerge from this limited variation calculation. The $\tilde{A}^{\nu(1)}$ are then used as approximations to the $\hat{A}^{(1)}$.

In support of this procedure let us first note that it yields a variational approximation to $\hat{E}^{(2)\prime}$. Proof: If we define $\tilde{E}^{\nu(2)\prime}$ by

$$\tilde{J}^{\nu(2)} = \tilde{E}^{\nu(2)\prime} (\hat{\Psi}^{(0)}, \hat{\Psi}^{(0)}) \qquad \text{(XXVI-12)}$$

then, since $\delta \hat{J}^{(2)} = 0$, all $\delta \hat{A}$, one can infer (or check by a detailed calculation using (XXV-14), (XXV-15) and (17) and (18) below) that

$$\overset{\vee}{J}^{(2)} = \hat{J}^{(2)} + O(\overset{\vee}{A}^{(1)} - \hat{A}^{(1)})^{\sim} \qquad \text{(XXVI-13)}$$

and therefore

$$\overset{\vee}{E}^{(2)\prime} = \hat{E}^{(2)\prime} + O(\overset{\vee}{A}^{(1)} - \hat{A}^{(1)})^{\sim} \qquad \text{(XXVI-14)}$$

which proves the point.

Further we will show that if $\hat{E}^{(0)}$ is itself a local minimum then $\overset{\vee}{E}^{(2)\prime}$ will be an upper bound to $\hat{E}^{(2)}$ so that in such cases the method is very well founded indeed.[1] We will give a very simple derivation of this result shortly; however first we will give a detailed proof, the excuse being our standard one that it will give us more practice with the machinery.

If we write

$$\overset{\vee}{\Psi}^{(1)} = \hat{\Psi}^{(1)} + \Delta^{(1)} \qquad \text{(XXVI-15)}$$

$$\overset{\vee}{\Psi}^{(2)\prime} = \hat{\Psi}^{(2)\prime} + \Delta^{(2)} \qquad \text{(XXVI-16)}$$

then from (9) and (XXV-3), and from (10) and (XXV-20) it follows that

$$\Delta^{(1)} = \frac{\partial \hat{\Psi}^{(0)}}{\partial \hat{A}^{(10)}} [\overset{\vee}{A}^{(1)} - \hat{A}^{(1)}] \qquad \text{(XXVI-17)}$$

and that

$$\Delta^{(2)} = \frac{1}{2} \frac{\partial^2 \hat{\Psi}^{(0)}}{\partial \hat{A}^{(10)\sim}} [\overset{\vee}{A}^{(1)\sim} - \hat{A}^{(1)\sim}] + \left(\frac{\partial^2 \hat{\Psi}}{\partial \nu \partial \hat{A}^{(00)}}\right)_0 [\overset{\vee}{A}^{(1)} - \hat{A}^{(1)}]$$

which can also be written as

$$\Delta^{(2)} = \frac{1}{2} \frac{\partial^2 \hat{\Psi}^{(0)}}{\partial \hat{A}^{(10)\sim}} (\overset{\vee}{A}^{(1)} - \hat{A}^{(1)})^{\sim} + \frac{\partial^2 \hat{\Psi}^{(10)}}{\partial \hat{A}^{(10)\sim}} \hat{A}^{(1)} (\overset{\vee}{A}^{(1)} - \hat{A}^{(1)}) + \left(\frac{\partial^2 \hat{\Psi}}{\partial \nu \partial \hat{A}^{(10)}}\right)_0 (\overset{\vee}{A}^{(1)} - \hat{A}^{(1)}) \qquad \text{(XXVI-18)}$$

However from (13) it follows that $(\check{J}^{(n)} - \hat{J}^{(n)})$ can contain no terms of first-order in $(\check{A}^{(n)} - \hat{A}^{(n)})$. Therefore we must have

$$\check{J}^{(n)} - \hat{J}^{(n)} = (\hat{\Psi}^{(0)}, (H^{(0)} - \hat{E}^{(0)})\Delta^{(n)'}) + (\Delta^{(n)'}, (H^{(0)} - \hat{E}^{(0)})\hat{\Psi}^{(0)}) + (\Delta^{(1)}, (H^{(0)} - \hat{E}^{(0)})\Delta^{(1)}) \tag{XXVI-19}$$

Where $\Delta^{(n)'}$ is the first term on the right in (18), i.e.,

$$\Delta^{(n)'} = \frac{1}{2} \frac{\partial^2 \hat{\Psi}^{(0)}}{\partial \hat{A}^{(0)} \partial \hat{A}^{(0)}} (\check{A}^{(n)} - \hat{A}^{(n)})^2$$

Comparison with (XXV-46), with $\delta \hat{A}$ in the latter taken to be

$$\delta \hat{A} = \delta\epsilon (\check{A}^{(n)} - \hat{A}^{(n)})$$

then yields the result that if $\delta^2 \hat{E}^{(0)} \neq 0$ then,

$$\check{J}^{(n)} \geq \hat{J}^{(n)}$$

and therefore

$$\check{E}^{(n)'} \geq \hat{E}^{(n)'} \tag{XXVI-20}$$

as claimed.

In the approach which we have followed, the functional $\tilde{J}^{(n)}$ did not emerge in a very natural way. However there is another way of proceeding, which we will now describe, in which $\tilde{J}^{(n)}$ appears quite automatically and in which its stationary (and minimal) properties are obvious. We first note that whether they are numerical parameters and/or functions, quite arbitrary $\tilde{A}$ (and hence $\hat{A}$) can be written as

$$\tilde{A} = \tilde{A}^{(0)} + \nu \tilde{A}^{(1)} + \nu^2 \tilde{A}^{(2)} + \cdots \tag{XXVI-21}$$

with the $\tilde{A}^{(n)}$, and their optimal values, independent of ν but otherwise quite arbitrary. Suppose however that we would now restrict these $\tilde{A}$ by requiring that $\tilde{A}^{(0)} = \hat{A}^{(0)}$. Evidently we would thereby be considering only a subset of the $\tilde{A}$ which we (implicitly) used earlier. However the important point is that since this is a subset which contains $\hat{A}$, it must lead to the same results.

Our new procedure then is to use as trial functions in the variation method, functions of the form

$$\tilde{\Phi} = \Theta(\tilde{A}, \nu) \tag{XXVI-22}$$

where

$$\tilde{A} = \hat{A}^{(0)} + \nu \tilde{A}^{(1)} + \nu^2 \tilde{A}^{(2)} + \cdots \tag{XXVI-23}$$

Expanding in powers of ν we then find

$$\tilde{\Phi} = \hat{\Phi}^{(0)} + \nu \tilde{\Phi}^{(1)} + \nu^2 \tilde{\Phi}^{(2)} + \cdots \tag{XXVI-24}$$

where $\tilde{\Phi}^{(1)}$ is as in (9), while

$$\tilde{\Phi}^{(1)} = \Phi^{(1)\prime} + \frac{\partial \hat{\Phi}^{(0)}}{\partial \hat{A}^{(0)}} \tilde{A}^{(1)} \tag{XXVI-25}$$

If now we insert these results into

$$(\tilde{\Phi}, (H - \tilde{E})\tilde{\Phi}) = 0 \tag{XXVI-26}$$

then we find

$$\tilde{E} = \hat{E}^{(0)} + \nu \tilde{E}^{(1)} + \nu^2 \tilde{E}^{(2)} + \cdots \tag{XXVI-27}$$

where

$$\left(\hat{\Psi}^{(1)}, (H^{(0)} - \hat{E}^{(0)})\hat{\Psi}^{(0)}\right) + \left(\hat{\Psi}^{(0)}, (H^{(0)} - \hat{E}^{(0)})\hat{\Psi}^{(1)}\right) + \left(\hat{\Psi}^{(0)}, (H^{(1)} - \hat{E}^{(1)})\hat{\Psi}^{(0)}\right) = 0 \tag{XXVI-28}$$

and (see below)

$$\hat{J}^{(1)} + \left(\hat{\Psi}^{(0)}, (H^{(2)} - \hat{E}^{(1)})\hat{\Psi}^{(0)}\right) = 0 \tag{XXVI-29}$$

etc.

We now require that

$$\delta\hat{E} = 0 \quad , \text{ all } \delta\hat{A}^{(1)}, \delta\hat{A}^{(2)} \cdots \tag{XXVI-30}$$

which will in turn imply the sequence of equations

$$\delta\hat{E}^{(n)} = 0 \qquad \text{all } \delta\hat{A}^{(1)}, \delta\hat{A}^{(2)} \cdots \delta\hat{A}^{(n)} \tag{XXVI-31}$$

In particular then we see from (29) that

$$\delta\hat{E}^{(1)} = 0 \qquad \text{all } \delta\hat{A}^{(1)}, \delta\hat{A}^{(2)} \tag{XXVI-32}$$

is equivalent to our earlier result that

$$\delta\hat{J}^{(1)} = 0 \qquad \text{all } \delta\hat{A}^{(1)}, \text{ all } \delta\hat{A}^{(2)} \tag{XXVI-33}$$

Also it immediately follows from (29) that the numerical value of $\hat{J}^{(1)}$ is $\hat{E}^{(1)}$, as we found before. Finally the minimal property of $\hat{J}^{(1)}$ when $\hat{E}^{(0)}$ is a local minimum follows easily since we can argue by continuity that if $\hat{E}^{(0)}$ is a local minimum for $\nu = 0$, it will remain a minimum for a small range of ν values around $\nu = 0$.

Therefore we will have

$$\tilde{E} \to \hat{E} \qquad \text{(XXVI-34)}$$

or

$$\hat{E}^{(0)} + \nu \hat{E}^{(1)} + \nu^2 \tilde{E}^{(2)}{}_{+-} \to \hat{E}^{(0)} + \nu \hat{E}^{(1)} + \nu^2 \hat{E}^{(2)}{}_{+-} \qquad \text{(XXVI-35)}$$

Cancelling the common terms on both sides, dividing by ν^2 and letting $\nu \to 0$ we then find

$$\tilde{E}^{(2)} \to \hat{E}^{(2)} \qquad \text{(XXVI-36)}$$

and hence, in particular

$$\tilde{E}^{(2)} \to \hat{E}^{(2)} \qquad \text{(XXVI-37)}$$

which, of course, is equivalent to

$$\tilde{E}^{(2)\prime} \to \hat{E}^{(2)\prime} \qquad \text{(XXVI-38)}$$

In writing down Eq. (35) and also earlier in writing down (29) we have used the fact that

$$\tilde{E}^{(0)} = \hat{E}^{(0)} \qquad \text{(XXVI-39)}$$

One can derive this by using (9) and (XXV-14) to show that the sum of the first two terms on the left in (28) vanishes. Alternatively, one can derive it completely within the present framework from $\delta \hat{E}^{(0)} = 0$ all $\delta \hat{A}^{(0)}$ by use of the particular choice $\delta \hat{A}^{(0)} = \delta\epsilon \hat{A}^{(1)}$. In a similar vein $\tilde{F}^{(2)}$ as it naturally appears in this approach, involves $\hat{F}^{(2)}$ and

therefore $\tilde{A}^{(2)}$ and $\hat{A}^{(1)}$, rather than $\tilde{\Psi}^{(2)}$ and therefore only $\hat{A}^{(1)}$. However, as we have discussed before, the two forms are quite equivalent, since the contributions of the $\tilde{A}^{(2)}$ terms to $\hat{J}^{(1)}$ and $\delta\hat{J}^{(1)}$ vanish identically.

What we have done in this chapter for $\hat{A}^{(1)}$ and $\hat{E}^{(1)}$ can be done for $\hat{A}^{(2)}$, $\hat{E}^{(4)}$ etc. Most simply, given $\hat{A}^{(0)}$ and $\hat{A}^{(1)}$ one can start again with $\tilde{A}$ of the form

$$\tilde{A} = \hat{A}^{(0)} + \nu\hat{A}^{(1)} + \nu^2\hat{A}^{(2)} + \cdots \qquad \text{(XXVI-40)}$$

and derive a functional which will yield a variational approximation to $\hat{E}^{(4)}$. Further since with (40), $\tilde{E}^{(3)}$ will equal $\hat{E}^{(3)}$ it follows, by the same argument which lead to (37), that this functional will yield an upper bound to $\hat{E}^{(4)}$ if $\hat{E}^{(0)}$ is a local minimum. Similarly given $\hat{A}^{(0)}$, $\hat{A}^{(1)}$, $\hat{A}^{(2)}$ one can start again and do the same for $\hat{A}^{(3)}$ and $\hat{E}^{(6)}$ etc.

Thus far, true to our original purpose, we have been discussing perturbation theory within a consistent variational calculation. However often one is confronted with rather a different situation. Namely one is simply given a function $\chi^{(0)}$ (derived perhaps from a variation calculation) as an approximation to an eigenfunction of $H^{(0)}$. The problem then is to perturb $\chi^{(0)}$ in some reasonable way by $\nu\Psi^{(1)} + \nu^2\Psi^{(2)} + \cdots$, in order to approximate the response of the actual system to the same perturbations. In such cases $H^{(0)}$ is usually the "accurate" Hamiltonian of an isolated system while $\nu\Psi^{(1)} + \nu^2\Psi^{(2)} + \cdots$ represent the effect of external fields and/or changes of parameters in $H^{(0)}$.

Several methods have been used to deal with such situations. One, of course, is simply to imbed $\chi^{(0)}$ in some consistent variation calculation of the type which we have been discussing. Another is to regard $\chi^{(0)}$ as defining a "model" system in a sense which we will discuss in Sec. XXXI. However here we wish to pursue a path which is suggested by the discussion following (21).

Namely suppose that we were to use in the variation method, trial functions of the form

$$\tilde{\Psi} = \chi^{(0)} + \nu \tilde{\chi}^{(1)} + \nu^2 \tilde{\chi}^{(2)} + \cdots \tag{XXVI-41}$$

where the $\tilde{\chi}^{(n)}$ contain various variational parameters and/or functions, and where we require that the $\tilde{\chi}^{(n)}$ and their optimal values be independent of ν. Then expanding

$$(\tilde{\Psi}, (H - \tilde{E})\tilde{\Psi}) = 0 \tag{XXVI-42}$$

we are led to a sequence of equations of which we will discuss only the first three:

$$(\chi^{(0)}, (H^{(0)} - \tilde{E}^{(0)})\chi^{(0)}) = 0 \tag{XXVI-43}$$

$$(\chi^{(0)}, (H^{(0)} - \tilde{E}^{(0)})\tilde{\chi}^{(1)}) + (\tilde{\chi}^{(1)}, (H^{(0)} - \tilde{E}^{(0)})\chi^{(0)}) + (\chi^{(0)}, (H^{(1)} - \tilde{E}^{(1)})\chi^{(0)}) = 0 \tag{XXVI-44}$$

$$(\chi^{(0)}, (H^{(0)} - \tilde{E}^{(0)})\tilde{\chi}^{(2)}) + (\tilde{\chi}^{(2)}, (H^{(0)} - \tilde{E}^{(0)})\chi^{(0)}) + (\tilde{\chi}^{(1)}, (H^{(0)} - \tilde{E}^{(0)})\tilde{\chi}^{(1)}) + (\tilde{\chi}^{(1)}, (H^{(1)} - \tilde{E}^{(1)})\chi^{(0)}) + (\chi^{(0)}, (H^{(1)} - \tilde{E}^{(1)})\tilde{\chi}^{(1)}) + (\chi^{(0)}, (H^{(2)} - \tilde{E}^{(2)})\chi^{(0)}) + \tag{XXVI-45}$$

etc.

Equation (43) just tells us that

$$\tilde{E}^{(0)} = (\chi^{(0)}, H^{(0)}\chi^{(0)})/(\chi^{(0)},\chi^{(0)}) \equiv \epsilon^{(0)} \qquad \text{(XXVI-46)}$$

More interesting is equation (44). Requiring

$$\delta\hat{E}^{(1)} = 0 \qquad \text{"all"}\ \delta\hat{\chi}^{(1)} \qquad \text{(XXVI-47)}$$

then yields

$$(\chi^{(0)}, (H^{(0)} - \epsilon^{(0)})\,\delta\hat{\chi}^{(1)}) + (\delta\hat{\chi}^{(1)}, (H^{(0)} - \epsilon^{(0)})\,\chi^{(0)}) = 0 \quad \text{"all"}\ \delta\hat{\chi}^{(1)} \qquad \text{(XXVI-48)}$$

and here we meet for the first time in all our discussion, the possibility of inconsistent equations. That is, thus far, since there were no obvious inconsistencies in the formalism, we have simply assumed that (at best in principle) all our equations had solutions. However we now encounter a situation where it is obvious that this need not be the case.

The analogue of equation (48) in our earlier discussion is the equation $\delta\hat{\epsilon}^{(1)} = 0$ all $\delta\hat{A}^{(1)}$, with $\tilde{E}^{(1)}$ given by (28). However that equation is automatically satisfied because of the nature of $\hat{\psi}^{(0)}$, i.e., it is just equation (XXV-14), and therefore, in particular, does not determine the $\hat{A}^{(1)}$ at all. Let us now consider (48). Suppose that as one extreme case $\hat{\chi}^{(1)}$ contains only linear parameters. Then (48) clearly does not determine those parameters. However equally clearly it does impose conditions on the basis set ϕ_μ, say. Namely we must have

$$(\phi_\mu, (H^{(0)} - \epsilon^{(0)})\,\chi^{(0)}) = 0 \qquad \text{(XXVI-50)}$$

and therefore, if one has not taken care to ensure this from the outset, one will have an inconsistency at this point (of course if $\chi^{(0)}$ has derived from a variational calculation for $H^{(0)}$ using this same basis set then all is consistent). Often (50) will be satisfied "naturally" for reason of symmetry, i.e., the perturbation $H^{(1)}$ is such that one will naturally choose the Φ_k to be orthogonal to $\chi^{(0)}$ and to $H^{(0)}\chi^{(0)}$. However in any case we would emphasize that if one does have linear parameters, then, for consistency (50) must be satisfied.

Let us now consider the other extreme in which $\hat{\chi}^{(1)}$ contains only non-linear parameters. Then usually, except for reasons of symmetry or because the calculation really falls into the pattern of Sec. XXV, (48) will, at last in part, determine the optimal values of these parameters.

We now turn to (45) which; of course, is very similar in structure to (29), and let us assume that $\hat{E}^{(0)}$ has been determined from (44). Linear parameters in $\hat{\chi}^{(1)}$ can now be determined by requiring $\delta \hat{E}^{(2)} = 0$. However again there are potential inconsistencies with linear parameters in $\hat{\chi}^{(1)}$ if they are not also contained in $\hat{\chi}^{(0)}$ since varying then will evidently yield equations like (50) again. Also if some of the nonlinear parameters in $\hat{\chi}^{(0)}$ have already been determined from (48), requiring that $\delta \hat{E}^{(2)} = 0$ when one varies them may also yield inconsistencies. However, in any case, one can use some set of parameters to calculate an approximate second-order energy from (45).

The potential difficulties we have already encountered clearly grow and multiply if one proceeds to higher order. The main lesson that we

would point to in all of this is that if one tries to "piece together" a variation-perturbation calculation in this way then, if one wishes to be consistent, one must exercise caution. However we would also emphasize, that if, usually through symmetry considerations, one has satisfied (48) without really restricting $\hat{\chi}^{(n)}$, then the only caution one must use with (45) and $\delta\hat{E}^{(n)}=0$ "all" $\delta\hat{\Sigma}^{(n)}$, $\delta\hat{\Psi}^{(n)}$, is with respect to the linear parameters in $\hat{\Sigma}^{(n)}$. Also of course, it may be that although (48) is not satisfied identically, still the left hand side is sufficiently small for practical purposes. We will return to some of these points in Sec. XXXI.

A topic we have not yet discussed is the following: we know that, depending on the nature of the set of trial functions, the $\hat{\Psi}$, and therefore the $\hat{\Psi}^{(n)}$, may satisfy various theorems. A natural question is then to what extent $\Psi^{(n)}$, will satisfy theorems analogous to those satisfied by $\hat{\Psi}^{(n)}$? We will postpone a discussion of this until Sec. XXXV.

XXVII. THE ν^{2n+1} THEOREM: INTERCHANGE THEOREMS

As we have discussed in Sec. XXIII, if one knows an eigenfunction correctly through order ν^n then one can calculate the eigenvalue correctly through order ν^{2n+1}. We will now derive a similar theorem within the variation method.[1]

Consider the function $\mathring{\Psi}$ defined by

$$\mathring{\Psi} = \mathcal{G}(\mathring{A}, \nu) \qquad \text{(XXVII-1)}$$

where

$$\mathring{A} = \hat{A}^{(0)} + \nu \hat{A}^{(1)} + \cdots + \nu^n \hat{A}^{(n)} \qquad \text{(XXVII-2)}$$

If we now define $\mathring{\Delta}$ and $\mathring{E}$ by

$$\mathring{\psi} = \hat{\psi} + \mathring{\Delta}$$

and

$$\mathring{E} = (\mathring{\psi}, H\,\mathring{\psi}) / (\mathring{\psi}, \mathring{\psi}) \qquad \text{(XXVII-3)}$$

then one readily finds that

$$\mathring{E} - \hat{E} = \left\{ (\mathring{\Delta}, (H-\hat{E})\hat{\psi}) + (\hat{\psi}, (H-\hat{E})\mathring{\Delta}) + (\mathring{\Delta}, (H-\hat{E})\mathring{\Delta}) \right\} / (\mathring{\psi}, \mathring{\psi}) \qquad \text{(XXVII-4)}$$

To determine the order of magnitude of $\mathring{E} - \hat{E}$ we first note that from (2) and (3) it follows that $(\mathring{A} - \hat{A})$ is of order ν^{n+1} and that

$$\mathring{\Delta} = (\mathring{A} - \hat{A}) \frac{\partial \hat{\psi}}{\partial \hat{A}} + O(\nu^{2n+2}) \qquad \text{(XXVII-5)}$$

Since $\mathring{\Delta}$ is of order ν^{n+1}, the last term on the right hand side of (4) is certainly of order ν^{2n+2}. Further if we were to multiply the first term in (5) by $\delta\epsilon$ then it would be of the form $\delta\hat{\psi} = \frac{\partial \hat{\psi}}{\partial \hat{A}} \delta\hat{A}$ with $\delta\hat{A} = \delta\epsilon(\mathring{A} - \hat{A})$. Thus its contribution to the sum of the first two terms on the right in (5) will vanish identically, so that the sum of these two terms is also of order ν^{2n+2}.

These results can be summarized in the "ν^{2n+1} theorem". Most generally, with a view toward the case of several perturbations, the theorem is that

$$\mathring{E} - \hat{E} = O((\mathring{A} - \hat{A})^2)$$

More specifically it is the statement that if we know $\hat{A}$ correctly through order ν^n then we can calculate $\hat{E}$ correctly through order ν^{2n+1} by expanding $\overset{\circ}{\psi}$ and then $\overset{\circ}{E}$ through order ν^{2n+1}. To recover the theorem for eigenfunctions as a special case we need only identify $\hat{A}$ with ψ.

Consider first $n \geq 0$, so that

$$\overset{\circ}{\psi} = \Theta(\hat{A}^{(0)}, \nu) = \hat{\psi}^{(0)} + \nu \left(\frac{\partial \psi}{\partial \nu}\right)_0 + \cdots \qquad \text{(XXVII-6)}$$

Then, writing (4) as

$$(\overset{\circ}{\psi}, (H - \overset{\circ}{E})\psi) = 0 \qquad \text{(XXVII-7)}$$

and expanding in power of ν we (CAN) then determine $\hat{E}^{(0)}$ and $\hat{E}^{(1)}$ since from the ν^{2n+1} theorem

$$\overset{\circ}{E} = \hat{E}^{(0)} + \nu \hat{E}^{(2)} + \cdots \qquad \text{(XXVII-8)}$$

Doing this we obviously recover (XXV-16) and (XXV-22). Similarly if we go on to $n=1$ we find

$$\psi = \hat{\psi}^{(0)} + \nu \hat{\psi}^{(1)} + \nu^2 \hat{\psi}^{(2)'} + \nu^3 \hat{\psi}^{(3)'} + \cdots \qquad \text{(XXVII-9)}$$

where we will not write out $\hat{\psi}^{(3)'}$ in detail (it is the part of $\hat{\psi}^{(3)}$ which does not involve $\hat{A}^{(2)}$ and $\hat{A}^{(3)}$). Then from

$$\overset{\circ}{E} = \hat{E}^{(0)} + \nu \hat{E}^{(1)} + \nu^2 \hat{E}^{(2)} + \nu^3 \hat{E}^{(3)} + \cdots$$

one recovers (XXV-20) and gains a new formula for $\hat{E}^{(3)}$.A

$$\{(\hat{\Psi}^{(3)}, (H^{(0)} - \hat{E}^{(0)})\hat{\Psi}^{(0)}) + (\hat{\Psi}^{(2)}, (H^{(0)} - \hat{E}^{(0)})\hat{\Psi}^{(1)}) +$$

$$+ (\hat{\Psi}^{(2)}, (H^{(1)} - \hat{E}^{(1)})\hat{\Psi}^{(0)}) + (\hat{\Psi}^{(1)}, (H^{(2)} - \hat{E}^{(2)})\hat{\Psi}^{(0)}) + c.c.\} +$$

$$+ (\hat{\Psi}^{(1)}, (H^{(1)} - \hat{E}^{(1)})\hat{\Psi}^{(1)}) + (\hat{\Psi}^{(0)}, (H^{(3)} - \hat{E}^{(3)})\hat{\Psi}^{(0)}) = 0$$

(XXVII-10)

Similarly with $M = 2$ one derives formulas for $\hat{E}^{(4)}$ and $\hat{E}^{(5)}$, etc., etc.

In all of our discussions we have assumed that $\hat{A}$ can be expanded in integral powers of ν. Suppose however that half integral powers are needed (we will give an example in a moment). Then it would appear that to get accuracy through order ν^{2n+1}, we would have to include the term $\nu^{n+1/2}\hat{A}^{(n+1/2)}$ in the definition of $\mathring{A}$ (in addition, of course, to including the lower half integral terms). However in some cases accuracy through order ν^n will still suffice. Thus if we leave the definition of $\mathring{A}$ essentially unchanged then

$$\mathring{A} = \hat{A} + O(\nu^{n+1/2})$$

so that instead of (5) we have

$$\mathring{A} = (\mathring{A} - \hat{A})\frac{\partial \hat{\Psi}}{\partial \hat{A}} + O(\nu^{2n+1}) \qquad \text{(XXVII-11)}$$

As it stands, the first term in $\mathring{A}$ appears to be of order $\nu^{n+1/2}$. However if, as we continue to assume, $\hat{\Psi}$ and $\mathring{\Psi}$ can be expanded in integral powers of ν it clearly must be of order ν^{n+1}. Therefore the last term in (11) is still of order ν^{2n+2}. Further the

first term in $\overset{\circ}{\Delta}$ still makes no contribution to the sum of the first two terms in (4). Therefore we have the result that if in addition

$$(H - \hat{E})\hat{\Psi} = O(\nu) \qquad \text{(XXVII-12)}$$

then we will again have

$$\overset{\circ}{E} - \hat{E} = O(\nu^{2n+2}) \qquad \text{(XXVII-13)}$$

As a case in point consider the ground state of 2-electron atomic ions with $\nu H^{(1)}$ the electron-electron interaction and $\Psi = |\tilde{R}_1\alpha\,\tilde{R}_2\beta| + |\tilde{R}_2\alpha\,\tilde{R}_1\beta|$. Then one finds[2] that although $\hat{\Psi}$ can be expanded in integral powers of $\nu = 1/Z$, the $\hat{A}$, namely $\hat{R}_1$ and $\hat{R}_2$, involve $\sqrt{\nu}$. However $\hat{\Psi}$ becomes exact for $\nu \to 0$ $(Z \to \infty)$ and hence (13) is satisfied. Thus in this case one does find that knowing the orbitals through order ν^m allows one to calculate $\hat{E}$ through order ν^{2m+3}.[3]

We now return to the subject of interchange theorems which we introduced in Sec. XXV. We will assume that there is no degeneracy (or more generally no non-analyticity problems) so that $\hat{A}$ can be expanded in a double series in ν_1 and ν_2, and also we will assume that only integer powers appear. Thus we have

$$\hat{A} = \hat{A}^{(0)} + \nu_1 \hat{A}_1^{(1)} + \nu_2 \hat{A}_2^{(1)} + \nu_1^2 \hat{A}_1^{(2)} + 2\nu_1\nu_2 \hat{A}_{12}^{(2)} + \cdots \qquad \text{(XXVII-14)}$$

We will now discuss various ways of calculating the energy correctly through order m_1 in ν_1 and order m_2 in ν_2. If

further the generalized Hellmann-Feynman theorem is satisfied for $\Gamma = \nu_1$,

$$\frac{\partial \hat{E}}{\partial \nu_1} = (\Psi, \frac{\partial H}{\partial \nu_1} \Psi)/(\Psi, \Psi) \qquad \text{(XXVII-15)}$$

then this will immediately enable us to also calculate the average value of $\partial H/\partial \nu_1$ through order $(m_1 - 1)$ in ν_1 and m_2 in ν_2 . However we will not assume that (15) is satisfied.

Let us first consider the case $m_2 \geq 1$. Then generalizing the direct approach of Sec. XXV, the ν^{2m+1} theorem implies that we should be able to calculate $\hat{E}$ through order ν^{m_1} by knowing $\hat{A}$ through order $[\ m_1\ ]$ where

$$[m_1] = m_1/2 \quad m_1 \text{ even}$$
$$= \frac{(m_1 - 1)}{2} \quad m_1 \text{ odd} \qquad \text{(XXVII-16)}$$

Indeed we see that we can get the desired results by using

$$\hat{A}^0 = \hat{A}^{(0)} + \nu_1 \hat{A}_1^{(1)} + \nu_1^2 \hat{A}_1^{(2)} + \cdots + \nu_1^{[m_1]} \hat{A}_1^{[m_1]} + \qquad \text{(XXVII-17)}$$
$$+ \nu_2 \hat{A}_2^{(1)} + 2\nu_2\nu_1 \hat{A}_{12}^{(1)} + \cdots + \nu_2\nu_1^{[m_1]} A_{12}^{([m_0]+1)}$$

Proof: From (14) and (16), the order of magnitude of $\hat{A}^0 - \hat{A}$ is

$$\hat{A}^0 - \hat{A} \sim \alpha \nu_1^{[m_1]+1} (1 + \beta\nu_2) + \gamma \nu_2^2 \qquad \text{(XXVII-18)}$$

where α , β , and γ are constant. Therefore from the ν^{2m+1} theorem

$$(\hat{E}^0 - \hat{E}) \sim [\alpha \nu_1^{[m_1]+1} (1 + \beta\nu_2) + \gamma \nu_2^2]^2$$

$$= \alpha^2 v_1^{2[n_1]+2}(1+\beta v_2)^2 + 2\alpha\gamma v_1^{[n_1]+1}(1+\beta v_2)v_2^2 + \gamma^2 v_2^4$$

(XXVII-19)

Then since from (17)

$$2[n_1]+2 = n_1+2 \quad n_1 \text{ even}$$
$$= n_1+1 \quad n_1 \text{ odd}$$

we see from (19) that $\overset{0}{E}$ will be accurate through terms of orders $v_1^{n_1}$ and v_2 as desired. In summary, in this direct approach we need $\hat{A}$ through order $[n_1]$ in v_1 and first-order in v_2, and the desired formula can be derived by expanding out $\overset{0}{E}$ to the requisite order. However we will not write out the details. In particular for $n_1=1$, we would need

$$\overset{0}{A} = \hat{A}^{(0)} + v_2 \hat{A}_2^{(1)} \qquad \text{(XXVII-20)}$$

while for $n_1=2$ we would use

$$\overset{0}{A} = \hat{A}^{(0)} + v_1 \hat{A}_1^{(1)} + v_2 \hat{A}_2^{(1)} + 2v_1v_2 \hat{A}_{12}^{(2)} \qquad \text{(XXVII-21)}$$

However the V^{2n+1} theorem also implies that we can calculate $\hat{E}$ through order v_2 by knowing $\hat{A}$ only for $v_2=0$, and we will now show, generalizing the discussion in Sec. XXV, that we get the desired result by using

$$\overset{0}{A} = \hat{A}^{(0)} + v_1 \hat{A}_1^{(1)} + \cdots v_1^{n_1} \hat{A}_1^{(n_1)} \qquad \text{(XXVII-22)}$$

instead of (17). Proof: Instead of (18) we have

$$\overset{0}{A} - \hat{A} \sim \alpha \nu_1^{m_1+1} + \beta \nu_2 \qquad \text{(XXVII-23)}$$

and therefore, from the ν^{2n+1} theorem

$$\overset{0}{E} - \hat{E} \sim (\alpha \nu_1^{m_1+1} + \beta \nu_2)^2 = \alpha^2 \nu_1^{2m_1+2} + 2\alpha\beta \nu_2 \nu_1^{m_1+1} + \beta^2 \nu_2^2 \qquad \text{(XXVII-24)}$$

which means that $\overset{0}{E}$ is accurate through terms of order $\nu_1^{m_1}$ and ν_2. Thus, in this interchanged form it suffices to calculate $\hat{A}$ through order m_1 in ν_1 and zero order in ν_2. In particular for $m_1 \geq 1$

$$\overset{0}{A} = \overset{(0)}{A} + \nu_1 \overset{(1)}{A_1} \qquad \text{(XXVII-25)}$$

and for $m_1 \geq 2$

$$\overset{0}{A} = \overset{(0)}{A} + \nu_1 \overset{(1)}{A_1} + \nu_1^2 \overset{(2)}{A_1} \qquad \text{(XXVII-26)}$$

Also, as in the example of Section XXV, the formulae which one derives by expanding $\overset{0}{E}$ take an especially simple form if the generalized Hellmann-Feynman theorem for $\sigma = 2\nu_2$ is satisfied.

The general case $m_2 \geq 1$ can be discussed in a similar way however since many more possibilities present themselves[4] we will not enter into details. Suffice it to say that the lowest order in ν_2 through which one needs to know $\hat{A}$ is of course $[\, m_2 \,]$ and this requires accuracy through order m_1 in ν_1, while the smallest order in ν_1 is $[\, m_1 \,]$ and this requires order m_2 in ν_2.

Also it should be kept in mind that whether or not (15) is satisfied and whether or not the double power series exists, we can obviously calculate

$$\left[\frac{(\hat{\Psi}, H_1^{(1)} \hat{\Psi})}{(\hat{\Psi}, \hat{\Psi})} \right]_{\nu_1=0}$$

accurately through order m_ν in ν_ν for any $H_1^{(1)}$ given $\hat{A}$ for $\nu_1=0$ accurately through order m_ν in ν_ν .

As an illustration of these results we will derive the interchanged formula for $\hat{E}^{(m)}_{1\nu}$ thereby generalizing (XXV-57). Using (25) we have

$$\hat{\Psi} = \Theta (\hat{A}^{(0)} + \nu \hat{A}_1^{(m)} ; \nu_1, \nu_\nu) \qquad \text{(XXVII-27)}$$
$$= \hat{\Psi}^{(0)} + \nu_1 \hat{\Psi}_1^{(m)} + \nu_\nu \hat{\Psi}_\nu^{(m)\prime} + 2\nu_1\nu_\nu \hat{\Psi}_{1\nu}^{(2m)\prime} + \cdots$$

where

$$\hat{\Psi}_1^{(m)} = \frac{\partial \hat{\Psi}^{(0)}}{\partial \hat{A}^{(0)}} \hat{A}_1^{(m)} + \left(\frac{\partial \hat{\Psi}}{\partial \nu_1} \right)_0 \qquad \text{(XXVII-28)}$$

$$\hat{\Psi}_\nu^{(m)\prime} = \left(\frac{\partial \hat{\Psi}}{\partial \nu_\nu} \right)_0 \qquad \text{(XXVII-29)}$$

and

$$\hat{\Psi}_{1\nu}^{(2m)\prime} = \frac{1}{2} \left(\frac{\partial^2 \hat{\Psi}}{\partial \hat{A}^{(0)} \partial \nu_\nu} \right) \hat{A}_1^{(0)} + \frac{1}{2} \left(\frac{\partial^2 \hat{\Psi}}{\partial \nu_1 \partial \nu_\nu} \right)_0 \qquad \text{(XXVII-30)}$$

where now the subscript 0 means $\nu_1 = \nu_\nu = 0$. Then picking out the term of order $\nu_1\nu_\nu$ in $\hat{E}^0$ from

$$(\hat{\Psi}^0, (H - \hat{E}^0) \hat{\Psi}^1) = 0$$

one finds, since from our theorem this term is $2\hat{E}^{(2m)}_{1\nu}$, that

$$2(\hat{\Psi}^{(0)}, (H^{(0)}_{12} - \hat{E}^{(0)}_{12})\hat{\Psi}^{(0)}) + \{2(\hat{\Psi}^{(0)}, (H^{(0)} - \hat{E}^{(0)})\hat{\Psi}^{(0)}_{12}{}')$$

$$+(\hat{\Psi}^{(1)}_{2}{}', (H^{(0)} - \hat{E}^{(0)})\hat{\Psi}^{(1)}_{1}) + (\hat{\Psi}^{(1)}_{2}{}', (H^{(1)}_{1} - \langle H^{(1)}_{1}\rangle_0)\hat{\Psi}^{(0)}) \qquad \text{(XXVII-31)}$$

$$+ (\hat{\Psi}^{(1)}_{1}, (H^{(1)}_{2} - \langle H^{(1)}_{2}\rangle_0)\hat{\Psi}^{(0)}) + c.c.\} = 0.$$

XXVIII. PERTURBATION ANALYSIS OF THE LINEAR VARIATION METHOD

Since the optimal functions of the linear variation method are exact eigenfunctions of the projected Hamiltonian $\overline{H}$ of Sec. VI one approach to perturbation analysis within the linear variation method is simply to apply Rayleigh Schrodinger perturbation theory to $\overline{H}$ written in the form

$$\overline{H} = \overline{H}^{(0)} + \nu\,\overline{H}^{(1)} + \nu^2\,\overline{H}^{(2)} + \cdots \qquad \text{(XXVIII-1)}$$

However this is a convenient approach only in case the set of trial functions is independent of ν in that $\overline{H}^{(n)} = \overline{H^{(n)}}$ and we can write

$$\overline{H} = \overline{H^{(0)}} + \nu\,\overline{H^{(1)}} + \nu^2\,\overline{H^{(2)}} + \cdots \qquad \text{(XXVIII-2)}$$

since otherwise one must, in addition to having the perturbation expansion of H, also have the perturbation expansion of Π. However if (2) applies one may now simply take over all the standard formulae, with the essential simplifying feature that since the $\overline{H^{(n)}}$ have no matrix elements outside the basis set, all sums in the sum over states formulae are finite and involve the matrix elements of the $\overline{H^{(n)}}$ between the $\hat{\Psi}^{(0)}_{L}$, i.e., the eigenfunctions of $\overline{H^{(0)}}$ within the basis

set. Further, within the basis set one may replace the $\overline{H^{(n)}}$ by the $H^{(n)}$ thus effecting an additional visual simplification. Indeed, particularly in the non-degenerate case, the perturbation series can be produced to arbitrarily high order by a simple iterative procedure.[1]

In order to carry out the procedure we have just sketched one must first produce the $\hat{\psi}_L^{(0)}$ and write everything in terms of them. However in some situations one may prefer not to do this, and would rather formulate things directly in terms of the original basis set. This is also the most convenient procedure for dealing with the general case (1).

In such situations then it is easiest to start from Eq. (V-12)

$$\sum_{L=1}^{M} (\phi_k, (H-\hat{E})\phi_L)\,\hat{a}_L = 0 \qquad k=1 \cdots M \qquad \text{(XXVIII-3)}$$

and develop the perturbation theory as the perturbation theory of this finite algebraic (i.e., finite matrix) eigenvalue problem, and where, in order to be quite general, we will allow the ϕ_k to depend on ν. It is now convenient to use a matrix notation. Thus we introduce the $M \times M$ Hermitian matrix $\mathcal{H}$ whose matrix elements are $(\phi_k, H \phi_L)$, the $M \times M$ Hermitian overlap matrix $\mathcal{S}$ whose matrix elements are (ϕ_k, ϕ_L), and the M-element column vector $\hat{\mathcal{a}}$ whose components are $\hat{a}_k$. Then (3) can be written compactly as

$$\mathcal{H}\hat{\mathcal{a}} = \hat{E}\,\mathcal{S}\,\hat{\mathcal{a}}$$

or, introducing another subscript to designate the state in which we are interested

$$\mathcal{H}\hat{\mathcal{a}}_k = \hat{E}_k \mathcal{S} \hat{\mathcal{a}}_k \qquad \text{(XXVIII-4)}$$

To carry out the perturbation analysis we now expand H and the basis set, and hence $\mathcal{H}$ and $\mathcal{S}$ in powers of ν :

$$\Phi_L = \Phi_L^{(0)} + \nu \Phi_L^{(1)} + \nu^2 \Phi_L^{(2)} + \cdots \qquad \text{(XXVIII-5)}$$

$$\mathcal{H} = \mathcal{H}^{(0)} + \nu \mathcal{H}^{(1)} + \nu^2 \mathcal{H}^{(2)} + \cdots \qquad \text{(XXVIII-6)}$$

$$\mathcal{S} = \mathcal{S}^{(0)} + \nu \mathcal{S}^{(1)} + \nu^2 \mathcal{S}^{(2)} + \cdots \qquad \text{(XXVIII-7)}$$

Then writing

$$\hat{Q}_K = \hat{Q}_K^{(0)} + \nu \hat{Q}_K^{(1)} + \nu^2 \hat{Q}_K^{(2)} + \cdots \qquad \text{(XXVIII-8)}$$

and

$$\hat{E}_K = \hat{E}_K^{(0)} + \nu \hat{E}_K^{(1)} + \nu^2 \hat{E}_K^{(2)} + \cdots \qquad \text{(XXVIII-9)}$$

we find the sequence of equations

$$(\mathcal{H}^{(0)} - \hat{E}_K^{(0)} \mathcal{S}^{(0)}) \hat{Q}_K^{(0)} = 0 \qquad \text{(XXVIII-10)}$$

$$(\mathcal{H}^{(0)} - \hat{E}_K^{(0)} \mathcal{S}^{(0)}) \hat{Q}_K^{(1)} + (\mathcal{H}^{(1)} - \hat{E}_K^{(1)} \mathcal{S}^{(0)} - \hat{E}_K^{(0)} \mathcal{S}^{(1)}) \hat{Q}_K^{(0)} = 0 \qquad \text{(XXVIII-11)}$$

$$(\mathcal{H}^{(0)} - \hat{E}_K^{(0)} \mathcal{S}^{(0)}) \hat{Q}_K^{(2)} + (\mathcal{H}^{(1)} - \hat{E}_K^{(1)} \mathcal{S}^{(0)} - \hat{E}_K^{(0)} \mathcal{S}^{(1)}) \hat{Q}_K^{(1)} + (\mathcal{H}^{(2)} - \hat{E}_K^{(2)} \mathcal{S}^{(0)} - \hat{E}_K^{(1)} \mathcal{S}^{(1)} - \hat{E}_K^{(0)} \mathcal{S}^{(2)}) \hat{Q}_K^{(0)} = 0 \qquad \text{(XXVIII-12)}$$

etc. These equations are of course precisely what one would get by applying (XXV-14), (XXV-15), etc. to this special case ($\hat{A} \to \hat{a}_L$). Thus for example, $\mathcal{H}^{(1)}$ includes contributions from $H^{(1)}$ combined with the $\Phi_L^{(0)}$ and from $H^{(0)}$ combined with the $\Phi_L^{(1)}$, the latter being the $(\partial \hat{\Phi}/\partial \nu)_0$ terms in (XXV-5), (XXV-11) and (XXV-15).

Also from (10) - (12) we have

$$\hat{E}^{(n)}_k = \{ \hat{Q}^{(0)}_k, (\mathcal{H}^{(n)} - \hat{E}^{(0)}_k \mathcal{S}^{(n)}) \hat{Q}^{(0)}_k \} / \{ \hat{Q}^{(0)}_k, \mathcal{S}^{(0)} \hat{Q}^{(0)}_k \} \qquad \text{(XXVIII-13)}$$

and

$$\hat{E}^{(2)}_k = \{ \hat{Q}^{(0)}_k, (\mathcal{H}^{(2)} - \hat{E}^{(1)}_k \mathcal{S}^{(1)} - \hat{E}^{(2)}_k \mathcal{S}^{(0)}) \hat{Q}^{(0)}_k \} / \{ \hat{Q}^{(0)}_k, \mathcal{S}^{(0)} \hat{Q}^{(0)}_k \} +$$

$$+ \{ \hat{Q}^{(0)}_k, (\mathcal{H}^{(2)} - \hat{E}^{(0)}_k \mathcal{S}^{(2)} - \hat{E}^{(1)}_k \mathcal{S}^{(1)}) \hat{Q}^{(1)}_k \} / \{ \hat{Q}^{(0)}_k, \mathcal{S}^{(0)} \hat{Q}^{(1)}_k \} \qquad \text{(XXVIII-14)}$$

etc. corresponding to (XXV-22) and (XXV-20) respectively. In (13) and (14) we have used the notation

$$\{ \mathcal{A}, \mathcal{B} \} = \sum_{k=1}^{M} a_k^* B_k \qquad \text{(XXVIII-15)}$$

Equation (10) is of course a standard matrix eigenvalue equation for which there are many methods of solution available. Let us assume that the equation has been solved, and let us further assume, for simplicity, that $\hat{E}^{(0)}_k$ is non-degenerate. Then with $\hat{E}^{(n)}_k$ computed from (13) in terms of zero order quantities, equation (11) yields a set of linear inhomogeneous equations to determine the components of $\hat{Q}^{(n)}_k$, and therefore $\hat{E}^{(n)}_k$ and $\hat{E}^{(n+1)}_k$. Again many solution methods are available. Here we will briefly sketch an approach which is especially useful in the case that the basis functions do not depend on ν so that

$$\mathcal{S}^{(1)} = \mathcal{S}^{(2)} = \cdots = 0 \qquad \text{(XXVIII-16)}$$

This method does involve a small rearrangement of the basis set.[2] Namely we replace one of the ϕ_k which occurs in $\hat{\Psi}^{(0)}_k$ by $\hat{\Psi}^{(0)}_k$ itself.

This clearly does not alter the set of trial functions. For notational convenience let us suppose that this ϕ_2 is ϕ_1. Then the only change in the matrices occurs in their first rows and first columns. In particular the matrix of $(\mathcal{H}^{(0)} - \hat{E}_k^{(0)} \mathcal{S}^{(0)})$ takes the form

$$\begin{bmatrix} 0 & 0 \cdots \\ 0 & \\ \vdots & \Omega \end{bmatrix} \qquad \text{(XXVIII-17)}$$

where Ω is an $(m-1) \times (m-1)$ matrix and is non singular if, as we will assume, $\hat{E}_k^{(0)}$ is non-degenerate. Also the 1-1 element of $(\mathcal{H}^{(0)} - \hat{E}_k^{(0)} \mathcal{S})$ becomes zero while only the first element of $\tilde{C}_k^{(0)}$ is different from zero, so that the vector $(\mathcal{H}^{(0)} - \hat{E}_k^{(0)} \mathcal{S}^{(0)}) \tilde{C}_k^{(0)}$ takes the form

$$\begin{pmatrix} 0 \\ V \end{pmatrix} \qquad \text{(XXVIII-18)}$$

where V is an $(M-1)$ element column vector.

Turning now to equation (10) we see that the first element of $\tilde{C}_k^{(1)}$ may now be chosen arbitrarily (and is physically irrelevant since, for example, it does not contribute to $\hat{E}_k^{(1)}$) while the remainder can be found from

$$-\Omega^{-1} V \qquad \text{(XXVIII-19)}$$

Further once we have $-\Omega^{-1}$ then we can go on and deal with the higher order equations, (12), etc., in similar fashion.

Returning to the general case, we can, for example, write the solution in a sum over states form. Thus since

$$(\mathcal{H}^{(0)} - \hat{E}_L^{(0)} S^{(0)})\,\hat{Q}_L^{(0)} = 0 \qquad \text{(XXVIII-20)}$$

it follows that

$$\{\hat{Q}_L^{(0)},\, S^{(0)}\, \hat{Q}_K^{(0)}\} = 0 \qquad E_K^{(0)} \neq E_L^{(0)} \qquad \text{(XXVIII-21)}$$

Therefore, by the usual arguments, we can choose the $\hat{Q}_L^{(0)}$ so that

$$\{\hat{Q}_L^{(0)},\, S^{(0)}\, \hat{Q}_K^{(0)}\} = \delta_{LK} \qquad \text{(XXVIII-22)}$$

With this choice of normalization (13) becomes

$$\hat{E}_K^{(1)} = \{\hat{Q}_K^{(0)},\, (\mathcal{H}^{(1)} - \hat{E}_K^{(0)} S^{(1)})\, \hat{Q}_K^{(0)}\} \qquad \text{(XXVIII-23)}$$

Further using (20) and (22) one readily finds that if $\hat{E}_K^{(0)}$ is non-degenerate then

$$\hat{Q}_K^{(1)} = -\sum_{L \neq K} \hat{Q}_L^{(0)}\, \frac{\{\hat{Q}_L^{(0)},\, (\mathcal{H}^{(1)} - \hat{E}_K^{(0)} S^{(1)})\, \hat{Q}_K^{(0)}\}}{\hat{E}_L^{(0)} - \hat{E}_K^{(0)}} \qquad \text{(XXVIII-24)}$$

in a solution of (11), which, with (14), then yields

$$\hat{E}_K^{(2)} = -\sum_{L \neq K} \frac{|\{\hat{Q}_K^{(0)},\, (\mathcal{H}^{(1)} - \hat{E}_K^{(0)} S^{(1)})\, \hat{Q}_L^{(0)}\}|^2}{\hat{E}_L^{(0)} - \hat{E}_K^{(0)}} +$$

$$+\{\hat{Q}_K^{(0)},\, (\mathcal{H}^{(2)} - \hat{E}_K^{(1)} S^{(1)} - \hat{E}_K^{(0)} S^{(2)})\, \hat{Q}_K^{(0)}\} \qquad \text{(XXVIII-25)}$$

and similarly for higher orders.

We will leave the discussion of the degenerate case to the interested reader.

For what follows it will be useful to generalize (23) and (25) to the case of multiple perturbations. Thus assuming that there are no degeneracy problems so that the multiple series for $\hat{E}$ and Ψ really exist, we may simply make the replacements

$$\mathcal{H}^{(n)} \longrightarrow v_1 \mathcal{H}^{(n)} + v_2 \mathcal{H}_2^{(1)} + \cdots \tag{XXVIII-26}$$

$$v^2 \mathcal{H}^{(n)} \longrightarrow v_1^2 \mathcal{H}_{11}^{(2)} + 2v_1 v_2 \mathcal{H}_{12}^{(2)} + \cdots \tag{XXVIII-27}$$

$$v S^{(n)} \longrightarrow v_1 S_1^{(n)} + v_2 S_2^{(1)} + \cdots \tag{XXVIII-28}$$

$$v^2 S^{(n)} \longrightarrow v_1^2 S_{11}^{(2)} + 2v_1 v_2 S_{12}^{(2)} + \cdots \tag{XXVIII-29}$$

to find

$$\hat{E}_k = \hat{E}_k^{(0)} + v_1 \hat{E}_{k1}^{(1)} + v_2 \hat{E}_{k2}^{(1)} + \cdots$$

$$+ v_1^2 \hat{E}_{k11}^{(2)} + 2v_1 v_2 \hat{E}_{k12}^{(2)} + \cdots + v_2^2 \hat{E}_{k22}^{(2)} + 2v_2 v_3 \hat{E}_{k23}^{(2)} + \cdots \tag{XXVIII-30}$$

where, from (23) and (25)

$$\hat{E}_{ks}^{(1)} = \{ \hat{Q}_k^{(0)}, (\mathcal{H}_s^{(1)} - \hat{E}_k^{(0)} S_s^{(1)}) \hat{Q}_k^{(0)} \} \tag{XXVIII-31}$$

and

$$\hat{E}_{kst}^{(2)} = -\frac{1}{2} \sum_{l \neq k} \Big[\{ \hat{Q}_k^{(0)}, (\mathcal{H}_s^{(1)} - \hat{E}_k^{(0)} S_s^{(1)}) \hat{Q}_l^{(0)} \} \{ \hat{Q}_l^{(0)}, (\mathcal{H}_t^{(1)} - \hat{E}_k^{(0)} S_t^{(1)}) \hat{Q}_k^{(0)} \}$$

$$+ (s \leftrightarrow t) \Big] / (\hat{E}_l^{(0)} - \hat{E}_k^{(0)}) +$$

$$+ \frac{1}{2} \{ \hat{Q}_k^{(0)}, (2\mathcal{H}_{st}^{(2)} - \hat{E}_s^{(1)} S_t^{(1)} - \hat{E}_t^{(1)} S_s^{(1)} - 2\hat{E}^{(0)} S_{st}^{(2)}) \hat{Q}_k^{(0)} \} \tag{XXVIII-32}$$

In recent years there has been much use made of perturbation theory within the linear variation method. In particular various $\hat{E}$ for two electron atomic ions have been expanded to very high orders in $1/Z$ [1,3,A] although in the original literature it has not always been made clear that there were expansions within the linear variation method. We will elaborate on this last rather cryptic comment, in Sec. XXX.

Thus far we have dealt with the "pure" linear variation method. That is we have assumed that the ϕ_k themselves contain no (non-linear) parameters; however, as we have mentioned on several occasions, in practice one often does include non-linear parameters $\tilde{\zeta}_\alpha$ in the ϕ_k. This of course will change the details of the analysis. Nevertheless as we saw in Sec. XXV - we can still determine the $\hat{A}^{(m)}$ and hence $\hat{E}_k^{(m)}$ and $\hat{\mathcal{E}}_k^{(m)}$ by solving a set of linear inhomogeneous equations, the equations which arise from (XXV-15). Recently Bishop and Macias[4], by proceeding in a rather different way, have shown how to write the resulting formula for $\hat{\mathcal{E}}_k^{(m)}$ in an interesting and useful form which makes extensive use of formulae like (32).

Following these authors then, let us start again and write

$$\hat{E}_k = \hat{E}_k(\hat{\zeta}, \nu) \tag{XXVIII-33}$$

where, in the explicit ν dependence we include the ν dependence of the linear variation parameters $\hat{a}_L$. Differentiating (22) then yields

$$\frac{d\hat{E}_k}{d\nu_s} = \sum_\alpha \frac{\partial \hat{E}_k}{\partial \hat{\zeta}_\alpha} \frac{\partial \hat{\zeta}_\alpha}{\partial \nu_s} + \frac{\partial \hat{E}_k}{\partial \nu_s} \tag{XXVIII-34}$$

$$= \frac{\partial \hat{E}_k}{\partial \nu_s}$$

since, from the variation method

$$\frac{\partial \hat{E}_\kappa}{\partial \hat{\xi}_\alpha} = 0 \qquad \text{(XXVIII-35)}$$

Then differentiating again we find

$$\frac{d^2 \hat{E}_\kappa}{d\nu_\tau \, d\nu_\sigma} = \sum_\alpha \frac{\partial^2 \hat{E}_\kappa}{\partial \hat{\xi}_\alpha \, \partial \nu_\sigma} \frac{\partial \hat{\xi}_\alpha}{\partial \nu_\tau} + \frac{\partial^2 \hat{E}_\kappa}{\partial \nu_\tau \, \partial \nu_\sigma} \qquad \text{(XXVIII-36)}$$

and therefore we have

$$\hat{E}^{(1)}_{\kappa\sigma} = \left(\frac{\partial \hat{E}_\kappa}{\partial \nu_\sigma} \right)_{\substack{\nu=0 \\ \xi=\xi_0}} \qquad \text{(XXVIII-37)}$$

$$\hat{E}^{(2)}_{\kappa\sigma} = \frac{1}{2} \sum_\alpha \left(\frac{\partial^2 \hat{E}_\kappa}{\partial \hat{\xi}_\alpha \, \partial \nu_\sigma} \right)_{\substack{\nu=0 \\ \hat{\xi}=\hat{\xi}_0}} \left(\frac{\partial \hat{\xi}_\alpha}{\partial \nu_\tau} \right)_{\substack{\nu=0 \\ \hat{\xi}=\hat{\xi}_0}} + \frac{1}{2} \left(\frac{\partial^2 \hat{E}_\kappa}{\partial \nu_\tau \partial \nu_\sigma} \right)_{\substack{\nu=0 \\ \hat{\xi}=\hat{\xi}_0}} \qquad \text{(XXVIII-38)}$$

where $\hat{\xi}_0$ denotes the values of the for .

Now $\hat{E}^{(1)}_{\kappa\sigma}$ and the last term on the right in (38), involve only differentiation with respect to the ν's with $\hat{\xi}$ fixed at $\hat{\xi}_0$. Therefore they are given by the right hand sides of (31) and (32) where all quantities are to be calculated assuming $\hat{\xi} = \hat{\xi}_0$ from the outset.

We will now show that the rest of $\hat{E}^{(2)}_{\kappa\sigma}$ can be calculated from similar formulae. First we observe from (35) that

$$\frac{d}{d\nu_\tau} \frac{\partial \hat{E}_\kappa}{\partial \hat{\xi}_\beta} = 0 \qquad \text{(XXVIII-39)}$$

and therefore

$$\sum_\alpha \frac{\partial^2 \hat{E}_\kappa}{\partial \hat{\xi}_\alpha \, \partial \hat{\xi}_\beta} \frac{\partial \hat{\xi}_\alpha}{\partial \nu_\tau} + \frac{\partial^2 \hat{E}_\kappa}{\partial \nu_\tau \, \partial \hat{\xi}_\beta} = 0 \qquad \text{(XXVIII-40)}$$

Thus putting $\nu=0$ and $\hat{\xi}=\hat{\xi}_0$ we see that the quantities

$$\left(\partial \hat{\xi}_\lambda / \partial \nu_\tau\right)_{\substack{\nu=0 \\ \hat{\xi}=\hat{\xi}_0}}$$

which we need in (38) can be calculated by solving linear inhomogeneous equations involving

$$\left(\partial^2 \hat{E}_\kappa / \partial \xi_\lambda \, \partial \xi_\beta\right)_{\substack{\nu=0 \\ \hat{\xi}=\hat{\xi}_0}} \tag{XXVIII-41}$$

and

$$\left(\partial^2 \hat{E}_\kappa / \partial \nu_\tau \, \partial \xi_\lambda\right)_{\substack{\nu=0 \\ \hat{\xi}=\hat{\xi}_0}} \tag{XXVIII-42}$$

the latter also being needed in (38) [incidentally note that if we multiply (40) by $\partial \xi_\beta / \partial \nu_\sigma$ and sum over β then it becomes clear that the first term on the right in (38) is really symmetric in σ and τ as it must be].

To complete the story we now note that (41) and (42) can all be calculated from formulae like (32) involving basis functions and derivatives of basis functions (in order to calculate the $\mathcal{H}^{(m)}$ and $\mathcal{J}^{(m)}$) evaluated at $\nu=0$, $\hat{\xi}=\hat{\xi}_0$. The point is that (32) is sufficiently general that it has become irrelevant whether the "perturbation parameters" in $\mathcal{H}$ and $\mathcal{J}$ occur both in H and the Φ_κ , or only in the Φ_κ . Thus by extending the set of ν_σ to include the

$$\mu_\alpha \equiv \hat{\xi}_\alpha - \hat{\xi}_{\alpha 0} \tag{XXVIII-43}$$

we can immediately apply the formulae since the second derivations that we want do treat the ν_σ and the μ_α as <u>independent</u> variables ($\partial / \partial \hat{\xi}_\alpha$

is clearly equivalent to $\partial/\partial\mu_\alpha$) and they are all to be evaluated at $\nu = \mu = 0$. For more details of this procedure and of its application to the calculation of force constants we refer the reader to reference 4 and to reference 2 of Sec. XXIV.

XXIX. PERTURBATION ANALYSIS OF UHF AND AUHF

Although the equations (IX-14) for the canonical spin orbitals of UHF

$$\hat{f}\,\hat{\varphi}_i = \epsilon_i\,\hat{\varphi}_i \qquad i = 1, \ldots N \tag{XXIX-1}$$

are nominally in Hamiltonian form, a more acurate description would be "pseudo-Hamiltonian" since $\hat{f}$ depends on the $\hat{\varphi}_j$, and also the equations (1) must be supplemented by the equations

$$(\hat{\varphi}_i, \hat{\varphi}_i) = 1 \qquad i = 1 \ldots N \tag{XXIX-2}$$

the equations themselves implying (for $\epsilon_i \neq \epsilon_j$.), or being consistent with

$$(\hat{\varphi}_i, \hat{\varphi}_j) = 0 \qquad i \neq j \tag{XXIX-3}$$

Let us first consider a situation in which the $H^{(n)}$ are one electron operators[A]

$$H^{(n)} = \sum_{s=1}^{N} h^{(n)}(s)$$

Then writing

$$\hat{\varphi}_i = \hat{\varphi}_i^{(0)} + \nu\,\hat{\varphi}_i^{(1)} + \cdots \tag{XXIX-4}$$

$$\epsilon_i = \epsilon_i^{(0)} + \nu\,\epsilon_i^{(1)} + \cdots \qquad \text{(XXIX-5)}$$

and using the explicit formula for $\hat{f}$ given in Eq. (IX-7) one finds the sequence of equations

$$(h^{(0)} - \epsilon_i^{(0)})\,\hat{\varphi}_i^{(0)} + \sum_j \left[(\hat{\varphi}_j^{(0)}, g\,\hat{\varphi}_j^{(0)})\,\hat{\varphi}_i^{(0)} - (\hat{\varphi}_j^{(0)}, g\,\hat{\varphi}_i^{(0)})\,\hat{\varphi}_j^{(0)} \right] = 0 \qquad \text{(XXIX-6)}$$

$$(\hat{\varphi}_k^{(0)}, \hat{\varphi}_k^{(0)}) = 1 \qquad \text{(XXIX-6')}$$

$$\begin{aligned}(h^{(0)} - \epsilon_i^{(0)})\,\hat{\varphi}_i^{(1)} &+ \sum_j \left[(\hat{\varphi}_j^{(0)}, g\,\hat{\varphi}_j^{(0)})\,\hat{\varphi}_i^{(1)} - (\hat{\varphi}_j^{(0)}, g\,\hat{\varphi}_i^{(1)})\,\hat{\varphi}_j^{(0)} \right] \\ &+ \sum_j \left[(\hat{\varphi}_j^{(0)}, g\,\hat{\varphi}_j^{(1)})\,\hat{\varphi}_i^{(0)} - (\hat{\varphi}_j^{(0)}, g\,\hat{\varphi}_i^{(0)})\,\hat{\varphi}_j^{(1)} \right] \\ &+ \sum_j \left[(\hat{\varphi}_j^{(1)}, g\,\hat{\varphi}_j^{(0)})\,\hat{\varphi}_i^{(0)} - (\hat{\varphi}_j^{(1)}, g\,\hat{\varphi}_i^{(0)})\,\hat{\varphi}_j^{(0)} \right] \\ &+ (h^{(1)} - \epsilon_i^{(1)})\,\hat{\varphi}_i^{(0)} = 0 \end{aligned} \qquad \text{(XXIX-7)}$$

$$(\hat{\varphi}_k^{(1)}, \hat{\varphi}_k^{(0)}) + (\hat{\varphi}_k^{(0)}, \hat{\varphi}_k^{(1)}) = 0 \qquad \text{(XXIX-7')}$$

etc. These same equations could, of course, be derived by applying (XXV-14), (XXV-15) etc. to this special case ($\hat{A} \to \hat{\varphi}_i$).

Since

$$\hat{\Psi}^{(0)} = |\hat{\varphi}_1^{(0)} \cdots \hat{\varphi}_N^{(0)}| \qquad \text{(XXIX-8)}$$

and

$$\hat{\Psi}^{(1)} = \sum_i |\hat{\varphi}_1^{(0)} \cdots \hat{\varphi}_i^{(1)} \cdots \hat{\varphi}_N^{(0)}| \qquad \text{(XXIX-9)}$$

where the $\hat{\varphi}_i^{(0)}$ are orthonormal we can use equations (3), (XXX-24) and (XXX-25) to calculate $\hat{E}^{(1)}$ and $\hat{E}^{(2)}$. The results, using Slater's rules, are

$$\hat{E}^{(1)} = \sum_i (\hat{\varphi}_i^{(0)}, h^{(1)} \hat{\varphi}_i^{(0)}) \qquad \text{(XXIX-10)}$$

and

$$\hat{E}^{(2)} = \sum_i (\hat{\varphi}_i^{(0)}, h^{(2)} \hat{\varphi}_i^{(0)}) + \frac{1}{2} \sum_i \{ (\hat{\varphi}_i^{(0)}, h^{(1)} \hat{\varphi}_{i\perp}^{(1)}) + c.c. \} \qquad \text{(XXIX-11)}$$

where $\hat{\varphi}_{i\perp}^{(1)}$ is the part of $\hat{\varphi}_i^{(1)}$ orthogonal to all the $\hat{\varphi}_k^{(0)}$:

$$\hat{\varphi}_{i\perp}^{(1)} = \hat{\varphi}_i^{(1)} - \sum_k \hat{\varphi}_k^{(0)} (\hat{\varphi}_k^{(0)}, \hat{\varphi}_i^{(1)}) \qquad \text{(XXIX-12)}$$

Note also that replacing $\hat{\varphi}_i^{(1)}$ by $\hat{\varphi}_{i\perp}^{(1)}$ in $\hat{\Psi}^{(1)}$ only changes $\hat{\Psi}^{(1)}$ by a pure imaginary (from (20')) multiple of $\hat{\Psi}^{(0)}$ namely

$$\left[-\sum_i (\hat{\varphi}_i^{(0)}, \hat{\varphi}_i^{(1)}) \right] \hat{\Psi}^{(0)}$$

and $\hat{\Psi}_\perp^{(1)}$ is therefore physically equivalent to $\hat{\Psi}^{(1)}$. Indeed, quite analogously to the situation in Rayleigh-Schrodinger perturbation theory, beyond the requirement (20') the $(\hat{\varphi}_i^{(0)}, \hat{\varphi}_i^{(1)})$ are undetermined by our equations and may be freely set equal to zero, essentially because our equations determine the $\hat{\varphi}_i$ and hence $\hat{\Psi}$ only up to arbitrary phase factors.

To calculate $\hat{E}^{(3)}$ from (XXVII-10) we also need

$$\hat{\Psi}^{(1)\prime} = \sum_{i \neq j} \sum | \hat{\varphi}_1^{(0)} \cdots \hat{\varphi}_i^{(1)} \cdots \hat{\varphi}_j^{(1)} \cdots \hat{\varphi}_N^{(0)} | \qquad \text{(XXIX-13)}$$

and

$$\hat{\Psi}^{(2)\prime} = \sum\sum\sum_{i \neq j \neq k} | \hat{\varphi}_1^{(0)} \cdots \hat{\varphi}_i^{(2)} \cdots \hat{\varphi}_j^{(1)} \cdots \hat{\varphi}_k^{(1)} \cdots \hat{\varphi}_N^{(0)} |$$

however we will not write out the detailed formula (which also involve only $\hat{\varphi}_{i\perp}^{(1)}$). Further we note that in accord with (3), the equations (7) imply

(for $\epsilon_i^{(0)} \neq \epsilon_j^{(0)}$) or are consistent[B] with

$$(\hat{\varphi}_i^{(1)}, \hat{\varphi}_j^{(0)}) + (\hat{\varphi}_i^{(0)}, \hat{\varphi}_j^{(1)}) = 0 \qquad i \neq j \tag{XXIX-14}$$

Using this result together with (7') and (12) it is then easy to see that the formula (11) for $\hat{E}^{(2)}$ can be rewritten as

$$\hat{E}^{(2)} = \sum_i (\hat{\varphi}_i^{(0)}, h^{(2)} \hat{\varphi}_i^{(0)}) + \frac{1}{2} \sum_i \{ (\hat{\varphi}_i^{(0)}, h^{(1)} \hat{\varphi}_i^{(1)}) + c.c. \} \tag{XXIX-15}$$

Equations (6) and (6') are standard UHF type equations and we will assume that they have been solved. Let us now consider equations (7) and (7'). As they stand they look rather forbidding. A natural way to solve them is to use an iterative procedure of the same type one would use to solve (6) and (6'). Namely let us write (6) and (7) in obvious notation as

$$(\hat{f}^{(0)} - \epsilon_i^{(0)}) \hat{\varphi}_i^{(0)} = 0 \tag{XXIX-16}$$

$$(\hat{f}^{(0)} - \epsilon_i^{(0)}) \hat{\varphi}_i^{(1)} + (\hat{f}^{(1)} + h^{(1)} - \epsilon_i^{(1)}) \hat{\varphi}_i^{(0)} = 0. \tag{XXIX-17}$$

Then one first "guesses" a set of spin orbitals $\overset{0}{\varphi}{}_k^{(1)}$ which satisfy

$$(\overset{0}{\varphi}{}_k^{(1)}, \hat{\varphi}_l^{(0)}) + (\hat{\varphi}_k^{(0)}, \overset{0}{\varphi}{}_l^{(1)}) = 0 \tag{XXIX-18}$$

so that, in accord with the UHF assumption the spin orbitals

$$\overset{0}{\varphi}_k = \hat{\varphi}_k^{(0)} + \nu \overset{0}{\varphi}_k^{(1)} \tag{XXIX-19}$$

are orthonormal through first-order in ν .

With these spin orbitals one calculates an $\overset{0}{f}{}^{(1)}$ and solves the more familiar looking uncoupled equations

$$(\hat{f}^{(0)} - \epsilon_i^{(0)})\,\varphi_i^{(1)} + (f^{0(1)} + h^{(1)} - \epsilon_i^{(1)})\,\hat{\varphi}_i^{(0)} = 0 \qquad \text{(XXIX-20)}$$

$$(\varphi_i^{(1)}, \hat{\varphi}_i^{(0)}) + (\hat{\varphi}_i^{(0)}, \varphi_i^{(1)}) = 0 \qquad \text{(XXIX-20')}$$

Further using (7) and the orthonormality of the $\{\varphi_k^{(0)}\}$ one readily finds that as a consequence of (19)

$$\epsilon_i^{(1)} = (\hat{\varphi}_i^{(0)}, (f^{0(1)} + h^{(1)})\,\hat{\varphi}_i^{(0)}) \qquad \text{(XXIX-21)}$$

and that [B]

$$(\varphi_i^{(1)}, \hat{\varphi}_k^{(0)}) + (\hat{\varphi}_i^{(0)}, \varphi_k^{(1)}) = 0 \qquad \text{(XXIX-22)}$$

finally the value of $(\hat{\varphi}_i^{(0)}, \varphi_i^{(1)})$ is not determined by (20) and hence may be chosen so as to satisfy (20'). For example it may be put equal to zero. Given the $\varphi_i^{(1)}$ one can then go on, calculate an $f^{(1)}$ and repeat the process, etc.

As we have seen, all one really needs are the $\hat{\varphi}_{i\perp}^{(1)}$. Hence it is of interest to note that the $\varphi_j^{(1)}$ in $\hat{f}^{(1)}$ can be replaced by $\hat{\varphi}_{j\perp}^{(1)}$ so that, by taking the projection of (17) orthogonal to the $\hat{\varphi}_k^{(0)}$ one can solve for the $\hat{\varphi}_{i\perp}^{(1)}$ separately (and then return, if desired, to calculate $\hat{\varphi}_{i\parallel}^{(1)}$). The point is simply that if in $\hat{f}^{(1)}$ one replaces $\hat{\varphi}_j^{(1)}$ by

$$\hat{\varphi}_{j\parallel}^{(1)} = \sum_k \hat{\varphi}_k^{(0)} (\hat{\varphi}_k^{(0)}, \hat{\varphi}_j^{(1)})$$

then, for example, the two direct terms in (7) become

$$D = \sum_j \sum_k (\hat{\varphi}_j^{(0)}, g\,\hat{\varphi}_k^{(0)})(\hat{\varphi}_k^{(0)}, \hat{\varphi}_j^{(1)})\,\hat{\varphi}_i^{(0)} +$$

$$+ \sum_j \sum_k (\hat{\varphi}_j^{(1)}, \hat{\varphi}_k^{(0)})(\hat{\varphi}_k^{(0)}, g\,\hat{\varphi}_j^{(0)})\,\hat{\varphi}_i^{(0)}$$

Then, interchanging the labels j and k in the second double sum and using (7') and (14), we see that $D = 0$; and in a similar fashion one can show that the contribution of the $\hat{\varphi}^{(1)}_{ji}$ to the exchange term also vanishes. Also, as a corollary, it evidently follows that we need only know $\hat{\varphi}^{(1)}_{j\perp}$ in order to calculate $\hat{f}^{(1)}$.

Thus far we have not worried about possible degeneracies among the $\epsilon_i^{(0)}$. Let us (NOW) suppose however that

$$\epsilon_k^{(0)} = \epsilon_i^{(0)} \qquad k \neq i,$$

then in the usual way (17) requires that the correct zero order spin orbitals satisfy

$$\left(\hat{\varphi}_k^{(0)}, \left(\hat{f}^{(1)} + h^{(1)}\right)\hat{\varphi}_i^{(0)}\right) = 0 \qquad \text{(XXIX-23)}$$

However we will now show that these requirements can be ignored in the calculation of the $\hat{\varphi}^{(1)}_{i\perp}$. To see this we first note that, in general there will be no problem (no vanishing energy denominator) in solving for the $\hat{\varphi}^{(1)}_{i\perp}$ even if we use the "wrong" $\hat{\varphi}^{(0)}_j$. Further since the right $\hat{\varphi}^{(0)}_j$ will be related to the wrong ones by some unitary transformation, one readily sees that since $\hat{f}^{(0)}$ and $\hat{f}^{(1)}$ are invariant to a common unitary transformation of the $\hat{\varphi}_k^{(0)}$ and of the $\hat{\varphi}_j^{(1)}$, that the "right" $\hat{\varphi}^{(1)}_{i\perp}$ will be the same unitary transformation of the "wrong" ones. However since $\hat{\Phi}^{(0)}$ and $\hat{\Phi}^{(1)}$ are also invariant, up to a common overall phase factor, to such simultaneous transformations it follows that we need make no special choice among degenerate $\hat{\varphi}_k^{(0)}$ when we calculate the $\hat{\varphi}^{(1)}_{i\perp}$. Of course if one for some reason (WISHES) to determine the correct $\hat{\varphi}_i^{(0)}$ and hence the correct $\hat{\varphi}_i^{(1)}$ one

must determine the unitary transformation U such a way as to satisfy (43). Since, as we have discussed, $\hat{f}^{(1)}$ is invariant to the unitary transformation, and hence can be computed from the wrong solutions, this is just the standard problem of diagonalizing an operator within a set of functions and we will not pursue it further.

We will now comment briefly on the possibility of a formal solution to (17) in a "sum over states" form. One approach would be to treat $(\hat{f}^{(1)} + h^{(1)} - \epsilon_i^{(0)})\,\varphi_i^{(0)}$ as an inhomogenity and expand in the eigenfunctions of $\hat{f}^{(0)}$ (N of which are the $\hat{\varphi}_k^{(0)}$). However this clearly is even more of a formal solution than usual because $\hat{f}^{(1)}$ is still unknown. However it can provide convenient starting point for the iteration process which we have sketched; or more correctly, for the AUHF analogue of that process.

Thus use of the eigenfunction of $\hat{f}^{(0)}$ does not immediately lead to a solution. Rather what one must do, and here we will not enter into detail but simply refer to the relevant literature, is to adopt the procedures which have been developed in UHF time-dependent perturbation theory, particularly by Thouless[2]; the adaptation consisting mainly in putting "ω" equal to zero.

The essential complication in equation (7) comes of course from the electron-electron interaction which produces the coupling among the $\hat{\varphi}_k^{(1)}$. A Specializing to atoms, if one is content with first-order accuracy in $1/Z$ in calculating $\hat{E}^{(1)}$, $\hat{E}^{(1)}$, etc then we know from our discussion of interchange theorems in Sec. XXVII that if $\hat{E}^{(0)}$ is non-degenerate it will suffice to solve the simpler equations which one gets

by simply deleting the electron-electron interaction, i.e., the g terms. Without the g terms Eq. (6), (7), etc. are of course simply the equations describing the perturbation of an electron in a coulomb field by $h^{(1)}$, $h^{(2)}$ etc. and for many interesting perturbations, can be solved analytically; the solution of the first N equations being sufficient, according to the discussion in Sec. V, to yield $\hat{E}$ correct through first order in $1/Z$ and N^{th} order in V. Also, as we discussed in Sec. XI, barring complications due to degeneracy in the exact problem, the resultant approximations to the $\hat{E}^{(n)}$ will agree with the exact results through this order in $1/Z$. Further since in UHF all generalized Hellmann-Feynman theorems are satisfied, one may use the simple formulae (XXV-57), (XXV-60) etc.

For details of such calculations we refer the reader to a series of papers by A. Dalgarno and colaborators, especially M. Cohen, starting in 1961 in the Proc. Roy. Soc. In these papers a variant of the procedures which we have described is also used, which is quite effective in increasing the numerical accuracy of the results. This method is called the Shielding Approximation[1] and consists in also deleting a part of the electron-nucleus interaction in (6), (7) etc. so that one is solving the problem of an electron in a coulombic field of charge $Z - S$ perturbed by $h^{(1)}$, $h^{(2)}$ etc. This problem is of course as readily soluble for $S \neq 0$ as for $S = 0$ and one then determines the value of S, the shielding constant, in such a way that the first-order corrections to these results vanish; $H_2^{(1)}$ of Sec. XXV now being the sum of the electron-electron interaction and the rest of the electron-nucleus interaction. We will say more about the shielding approximation in Sec. XXXVI.

If equations (7) and (7') cannot be handled analytically or numerically, a common way to approximate their solution is to use a variation method. From our discussions in XXVI, especially that following Eq. (XXVI-23) it is clear that a simple way to derive the relevant functional is to go back to our formula (IX-5) for $\tilde{\Xi}$ in UHF, insert

$$\tilde{\varphi}_i = \hat{\varphi}_i^{(0)} + \lambda \tilde{\varphi}_i^{(1)} + \cdots \qquad \text{(XXIX-24)}$$

and extract the second order terms. In the process terms involving $\tilde{\varphi}_i^{(2)}$ seem to be present, however in accord with our general considerations, they can be eliminated in favor of terms involving $\tilde{\varphi}_i^{(1)}$. Namely one finds that they occur in the combination

$$(\hat{\varphi}_i^{(2)}, \hat{f}^{(0)} \hat{\varphi}_i^{(0)}) + (\hat{\varphi}_i^{(0)}, \hat{f}^{(0)} \tilde{\varphi}_i^{(2)}) \qquad \text{(XXIX-25)}$$

which from (7) can be written

$$\epsilon_i^{(0)} [(\hat{\varphi}_i^{(2)}, \hat{\varphi}_i^{(0)}) + (\hat{\varphi}_i^{(0)}, \tilde{\varphi}_i^{(2)})] \qquad \text{(XXIX-26)}$$

However the assumed orthonormality of the $\tilde{\varphi}_i$ requires that

$$(\tilde{\varphi}_i^{(1)}, \hat{\varphi}_j^{(0)}) + (\hat{\varphi}_i^{(0)}, \tilde{\varphi}_j^{(1)}) = 0 \qquad \text{(XXIX-27)}$$

and

$$(\tilde{\varphi}_i^{(2)}, \hat{\varphi}_j^{(0)}) + (\hat{\varphi}_j^{(0)}, \tilde{\varphi}_i^{(2)}) + (\tilde{\varphi}_i^{(1)}, \tilde{\varphi}_j^{(1)}) = 0 \qquad \text{(XXIX-28)}$$

so that we can replace (26) by

$$-\epsilon_i^{(0)} (\tilde{\varphi}_i^{(1)}, \tilde{\varphi}_i^{(1)}) \qquad \text{(XXIX-29)}$$

We will not write out the functional $\check{J}^{(n)}$ in detail but will only state that

$$\delta \check{J}^{(n)} = 0 \qquad \text{all } \delta \check{\varphi}_i^{(n)} \tag{XXIX-30}$$

then yields

$$\sum_i \left[(\delta \check{\varphi}_i^{(n)}, \check{X}_i) + (\check{X}_i, \delta \check{\varphi}_i^{(n)}) \right] = 0 \tag{XXIX-31}$$

where $\check{X}_i$ has the structure of the left hand side of (7) exclusive of the term $-\epsilon_i^{(n)} \hat{\varphi}_i^{(n)}$, but with the $\hat{\varphi}_k^{(n)}$ replaced by $\check{\varphi}_k^{(n)}$. Also we know, from the results of Sec. XXVI that $\check{J}^{(n)}$ yields an upper bound to $\hat{E}^{(n)\prime}$ if $\hat{E}^{(n)}$ is itself a local minimum.

The equations (31) are of course not the whole story. In addition we have the constraints (27) and hence

$$(\check{\varphi}_i^{(n)}, \hat{\varphi}_j^{(n)}) + (\hat{\varphi}_i^{(n)}, \check{\varphi}_j^{(n)}) = 0 \tag{XXIX-32}$$

Now a first reaction might be to incorporate these constraints by means of diagonal and off diagonal Lagrange multipliers. However clearly $\check{J}^{(n)}$ supplemented only by diagonal multipliers already yields the variational equivalent of (7). That is if we define $\hat{J}^{(n)\prime}$ by

$$\hat{J}^{(n)\prime} = \check{J}^{(n)} + \sum_i \epsilon_i^{(n)} \left[(\hat{\varphi}_i^{(n)}, \hat{\varphi}_i^{(n)}) + (\check{\varphi}_i^{(n)}, \hat{\varphi}_i^{(n)}) \right] \tag{XXIX-33}$$

Then from (31) we see that

$$\delta \hat{J}^{(n)\prime} = 0 \qquad \text{all } \delta \hat{\varphi}^{(n)} \tag{XXIX-34}$$

is equivalent to (7).

Thus if the $\delta\check{\varphi}_i$ could be completely arbitrary the off-diagonal Lagrange multipliers would certainly be unnecessary. We will now show that only a rather natural flexibility in the trial functions is already sufficient to ensure that (32) will be satisfied for $i \neq j$ without the use of off-diagonal Lagrange multipliers (of course it may be satisfied automatically by symmetry in which case all our maneuverings are unnecessary). Since the electron-electron interaction does not change the situation qualitatively, let us put $g = 0$ in order not to obscure the argument with long formulae.

Equations (31) supplemented by diagonal Lagrange multipliers are then in detail

$$\sum \{(\delta\check{\varphi}_i, (h^{(0)} - \epsilon_i^{(0)})\check{\varphi}_i^{(1)} + (h^{(1)} - \check{\epsilon}_i^{(1)})\hat{\varphi}_i^{(0)}) + c.c.\} = 0 \qquad \text{(XXIX-35)}$$

"or" $\delta\check{\varphi}_i^{(1)}$

where the $\check{\epsilon}_i^{(1)}$ are to be determined by the requirement

$$(\check{\varphi}_i^{(1)}, \hat{\varphi}_i^{(0)}) + (\hat{\varphi}_i^{(0)}, \check{\varphi}_i^{(1)}) = 0 \qquad \text{(XXIX-36)}$$

We will not show that if

$$\delta\check{\varphi}_i^{(1)} = \delta a\, \hat{\varphi}_k^{(0)} \qquad i,k = 1 \cdots N \qquad \text{(XXIX-37)}$$

with δa an infinitesimal but otherwise arbitrary complex number are possible variations of the $\check{\varphi}_i^{(1)}$ within the set of trial functions then we will have

$$(\check{\varphi}_i^{(1)}, \hat{\varphi}_j^{(0)}) + (\hat{\varphi}_i^{(0)}, \check{\varphi}_j^{(1)}) = 0 \qquad i \neq j \qquad \text{(XXIX-38)}$$

The simplest way to ensure that (37) are possible variations is of course to include linear variational parameters multiplying the $\hat{\varphi}_k^{(0)}$ in the

set of trial functions. The proof that (38) will then hold is simple. Choose

$$\delta \check{\varphi}_i^{(1)} = \delta c \, \hat{\varphi}_K^{(0)} \qquad i = L \qquad \text{(XXIX-39)}$$
$$= 0 \qquad i \neq L$$

for some K and L, with $K \neq L$. Then in the usual way (35) yields

$$(\hat{\varphi}_K^{(0)}, \; (h^{(0)} - \epsilon_L^{(0)}) \check{\varphi}_L^{(1)} + (h^{(1)} - \check{\epsilon}_L^{(1)}) \hat{\varphi}_L^{(0)}) = 0 \qquad \text{(XXIX-40)}$$

or,

$$(\epsilon_K^{(0)} - \epsilon_L^{(0)}) (\hat{\varphi}_K^{(0)}, \check{\varphi}_L^{(1)}) + (\hat{\varphi}_K^{(0)}, (h^{(1)} - \check{\epsilon}_L^{(1)}) \hat{\varphi}_L^{(0)}) = 0 \qquad \text{(XXIX-41)}$$

which, for $K = L$, tells us that

$$\check{\epsilon}_K^{(1)} = (\hat{\varphi}_K^{(0)}, \; h^{(1)} \hat{\varphi}_K^{(0)}) = \hat{\epsilon}_K^{(1)} \qquad \text{(XXIX-42)}$$

and leaves $(\hat{\varphi}_K^{(0)}, \check{\varphi}_K^{(1)})$ arbitrary. Similarly interchanging K and L and taking the complex conjugate we find

$$(\epsilon_L^{(0)} - \epsilon_K^{(0)}) (\check{\varphi}_K^{(1)}, \hat{\varphi}_L^{(0)}) + (\hat{\varphi}_K^{(0)}, (h^{(1)} - \check{\epsilon}_K^{(1)}) \hat{\varphi}_L^{(0)}) = 0$$

Subtracting this from (41) then yields for $K \neq L$

$$(\epsilon_L^{(0)} - \epsilon_K^{(0)}) \left[(\hat{\varphi}_K^{(0)}, \check{\varphi}_L^{(1)}) + (\check{\varphi}_K^{(1)}, \hat{\varphi}_L^{(0)}) \right] = 0 \qquad \text{(XXIX-43)}$$

which proves the point for $\epsilon_K^{(0)} \neq \epsilon_L^{(0)}$. If $\epsilon_K^{(0)} = \epsilon_L^{(0)}$, $K \neq L$ then this procedure also ensures that $(\hat{\varphi}_K^{(0)}, \check{\varphi}_L^{(1)})$ is left undetermined and hence can be assumed to satisfy (38).[B]

One final remark on the use of this variation method: Often it is used with the $\hat{\varphi}_i^{(0)}$ replaced by approximations, usually of the AUHF

type. If one uses only linear variation parameters in the $\hat{\varphi}_c^{(0)}$ and if the same basis set[C] is used as in $\hat{\varphi}_c^{(0)}$ (or if a subset of that basis set is used) then one can be said to be doing a consistent calculation within AUHF (or a variational approximation thereto). However otherwise the theoretical foundation is not so clear (we will discuss an analogous situation in Sec. XXXI). In any case one cannot say with confidence that $\tilde{E}^{(3)} \geq \hat{E}^{(2)}$, even if $\hat{E}^{(0)}$ is a local minimum.

We now turn briefly to the case of two-electron perturbations, and in particular to the $1/Z$ expansion for atoms in which the entire coulomb interaction between the electrons is treated as a perturbation. Thus instead of (6) and (6') we have[3]

$$(h - \epsilon_i^{(0)})\hat{\varphi}_i^{(0)} = 0 \qquad \text{(XXIX-44)}$$

$$(\hat{\varphi}_i^{(0)}, \hat{\varphi}_i^{(0)}) = 1 \qquad \text{(XXIX-44')}$$

and instead of (7) and (7')

$$(h - \epsilon_i^{(0)})\hat{\varphi}_i^{(1)} + \sum_j (\hat{\varphi}_j^{(0)}, g\hat{\varphi}_j^{(0)})\hat{\varphi}_i^{(0)} - (\hat{\varphi}_j^{(0)}, g\hat{\varphi}_i^{(0)})\hat{\varphi}_j^{(0)} = \epsilon_i^{(1)}\hat{\varphi}_i^{(0)} \qquad \text{(XXIX-45)}$$

$$(\hat{\varphi}_i^{(0)}, \hat{\varphi}_i^{(1)}) + (\hat{\varphi}_i^{(1)}, \hat{\varphi}_i^{(0)}) = 0 \qquad \text{(XXIX-45')}$$

Equations (45) are simpler than (7) in that there is no coupling between the $\hat{\varphi}_i^{(1)}$.

The solutions of (44) and (44') are of course hydrogenic functions, and Eq. (45) and (45') can then often be solved analytically[4] though, in any case, variational techniques are available. Thus we will not pursue the general discussion further except to call attention to an interesting

formula[5] which relates the $\epsilon_i^{(n)}$ to the $\hat{E}^{(n)}$.

From the virial theorem (which is satisfied in UHF)

$$2\hat{T} + \hat{V}_N + \hat{V}_e = 0 \qquad \text{(XXIX-46)}$$

or

$$2\hat{E} = \hat{V}_N + \hat{V}_e \qquad \text{(XXIX-47)}$$

where $\hat{V}_N$ is the average electron-nucleus potential energy and $\hat{V}_e$ is the average electron-electron potential energy. Further, from the generalized Hellmann-Feynman theorem for $\gamma = Z$ (also satisfied in UHF) we have

$$Z \frac{\partial \hat{E}}{\partial Z} = \hat{V}_N \qquad \text{(XXIX-48)}$$

and also we know that in UHF (Eq. (IX-27))

$$\hat{V}_e = \sum \epsilon_i - \hat{E} \qquad \text{(XXIX-49)}$$

Combining these results then we have

$$3\hat{E} - Z \frac{\partial \hat{E}}{\partial Z} = \sum \epsilon_i \qquad \text{(XXIX-50)}$$

If now we expand $\hat{E}$ and ϵ_i according to

$$\hat{E} = Z^2 \sum_n \hat{E}^{(n)} Z^{-n} \qquad \text{(XXIX-51)}$$

and

$$\epsilon_i = Z^2 \sum_n \epsilon_i^{(n)} Z^{-n} \qquad \text{(XXIX-52)}$$

then (50) immediately yields the desired formula

$$(N+1)\ \hat{E}^{(N)} = \sum_i \epsilon_i^{(N)} \qquad \text{(XXIX-53)}$$

The discussion of AUHF[5] parallels that of UHF (except for the appearance of extra terms if the functions in the basis set depend on) so that we need not enter into too many details. In particular for a one-electron perturbation[6] the formulae (10) and (11) can be simply taken over unless the basis set depends on ν in which case one must start from (XXV-22) and (XXV-20). Then instead of (6) and (7) we will have, adapting the notation of Sec. IX and also introducing the $M \times M$ matrix $h^{(1)}$ whose elements are $(\chi_\alpha^{(0)}, h^{(1)} \chi_\beta^{(0)})$

$$(\hat{f}^{(0)} - \epsilon_i^{(0)} S^{(0)})\, \hat{c}_i^{(0)} = 0 \qquad \text{(XXIX-54)}$$

and

$$(\hat{f}^{(0)} - \epsilon_i^{(0)} S^{(0)})\, \hat{c}_i^{(1)} + (\hat{f}^{(1)} + h^{(1)} - \epsilon_i^{(1)} S^{(0)} - \epsilon_i^{(0)} S^{(1)})\, \hat{c}_i^{(0)} = 0 \qquad \text{(XXIX-55)}$$

where $\hat{f}^{(1)}$ has one piece analogous to $\hat{f}^{(1)}$ and another which arises from any ν dependence of the functions in the basis set - the latter piece then involving only the $\hat{c}_j^{(0)}$ and not the $\hat{c}_j^{(1)}$. In more detail these contributions are

$$\sum_j \left\{ \frac{\partial \hat{f}^{(0)}}{\partial \hat{c}_j^{(0)}}\, \hat{c}_j^{(0)} + \frac{\partial \hat{f}^{(0)}}{\partial \hat{c}_j^{(0)*}}\, \hat{c}_j^{(0)*} \right\} \qquad \text{(XXIX-56)}$$

and

$$\sum_\alpha \left\{ \frac{\partial \hat{f}^{(0)}}{\partial \chi_\alpha^{(0)}} \left(\frac{\partial \chi_\alpha}{\partial \nu}\right)_0 + \frac{\partial \hat{f}^{(0)}}{\partial \chi_\alpha^{(0)*}} \left(\frac{\partial \chi_\alpha^*}{\partial \nu}\right)_0 \right\} \qquad \text{(XXIX-57)}$$

respectively.

In addition, equations (54) and (55) are supplemented by conditions analogous to (6') and (7'), these conditions being

$$\{ \hat{\mathcal{C}}_i^{(0)}, S^{(0)} \hat{\mathcal{C}}_i^{(0)} \} = 1 \qquad \text{(XXIX-58)}$$

and

$$\{ \hat{\mathcal{C}}_i^{(1)}, S^{(0)} \hat{\mathcal{C}}_i^{(0)} \} + \{ \hat{\mathcal{C}}_i^{(0)}, S^{(0)} \hat{\mathcal{C}}_i^{(1)} \} + \{ \hat{\mathcal{C}}_i^{(0)}, S^{(1)} \hat{\mathcal{C}}_i^{(0)} \} = 0 \qquad \text{(XXIX-59)}$$

The equations themselves then imply, or are consistent with

$$\{ \hat{\mathcal{C}}_i^{(0)}, S^{(0)} \hat{\mathcal{C}}_j^{(0)} \} = 0 \qquad i \neq j \qquad \text{(XXIX-60)}$$

and

$$\{ \hat{\mathcal{C}}_i^{(1)}, S^{(0)} \hat{\mathcal{C}}_j^{(0)} \} + \{ \hat{\mathcal{C}}_i^{(0)}, S^{(0)} \hat{\mathcal{C}}_j^{(1)} \} + \{ \hat{\mathcal{C}}_i^{(0)}, S^{(1)} \hat{\mathcal{C}}_j^{(0)} \} = 0 \qquad i \neq j \qquad \text{(XXIX-61)}$$

Given the solutions to (54) and (58), equations (55) and (59) then form a set of linear, inhomogeneous, coupled equations for the components of the $\hat{\mathcal{C}}_i^{(1)}$ and the $\epsilon_i^{(1)}$, and such equations can be solved in various ways, for example by iteration on the coupling terms. The algebra is greatly simplified if one expands the $\hat{\mathcal{C}}_i^{(1)}$ in terms of the M solutions of

$$(\hat{\mathcal{F}}^{(0)} - \epsilon_\alpha^{(0)} S^{(0)}) B_\alpha^{(0)} = 0 \qquad \alpha = 1 \ldots M \qquad \text{(XXIX-62)}$$

the $B_\alpha^{(0)}$ being either automatically, or by choice, orthonormal:

$$\{ B_\alpha^{(0)}, S^{(0)} B_\beta^{(0)} \} = \delta_{\alpha\beta}$$

The $\hat{\mathcal{C}}_i^{(0)}$ of course comprise N of these solutions - the "occupied" $B_\alpha^{(0)}$.

Then, analogous to the situation in UHF, one readily sees that one can solve separately for the $(M-N)\cdot N$ coefficients of the "unoccupied" $B_\alpha^{(0)}$ [Because of the $S^{(1)}$ term in (59), the coefficients of the occupied $B_\alpha^{(0)}$ don't make a zero contribution to $\hat{f}^{(1)}$, however it is a contribution which can be expressed in terms of the zero order quantities and $S^{(1)}$]. Also, just as in UHF, one can show that these coefficients are sufficient to determine $\hat{E}^{(1)}$, $\hat{E}^{(2)}$ and $\hat{E}^{(3)}$ and that one can ignore any ambiguities due to degeneracy as long as one does not want to calculate the coefficients of the occupied $B_{s\alpha}^{(0)}$.

For OHF and AUHF[7] the equations are further complicated by the presence of off diagonal Lagrange multipliers $\epsilon_{st}^{(m)}$. However in at least two cases of interest they usually cause no real trouble. First, if one neglects the electron-electron interactions completely as in the shielding approximation for one-electron perturbations, one usually can choose

$$\epsilon_{st}^{(0)} = 0 \qquad s \neq t$$

Thus in the m^{th} order equations the $\epsilon_{st}^{(0)}$ cause no coupling of the $\psi_i^{(m)}$. Further the $\epsilon_{st}^{(m)}$ will be part of the inhomogenity so that the solutions are linear in the $\epsilon_{st}^{(m)}$ and hence the latter can (in principle at least) be readily determined by imposing the orthogonality constraints. Of course if the $h^{(m)}$ are all spin independent then in this approximation one can choose $\epsilon_{st}^{(m)} = 0 \quad s \neq t$ for all $s \neq t$ since then in the $Z \to \infty$ limit all the $\hat{\psi}_i$ will obviously be eigenfunctions of

$$h^{(0)} + \nu h^{(1)} + \cdots$$

The other case occurs in the $1/Z$ expansion of the energy of an isolated atom. Here again one can usually choose

$$\epsilon^{(0)}_{st} = 0 \qquad s \neq t$$

and similar considerations apply. For more details on these special cases, and on the general situation, we refer the interested reader to the literature.[4,8]

XXX. PERTURBATION ANALYSIS OF THE VARIATION PRINCIPLE - GENERAL FORMALISM

By identifying $\hat{A}$ with ψ the perturbation analysis of the variation principle can be derived from the results of Sec. XXV and XXVI. However in the interests of clarity we will pretty much start all over; though the parallels with the previous discussions should be obvious.

By use of Rayleigh Schroedinger perturbation theory one can determine the eigenfunctions and eigenvalues of H as a power series in ν. Thus

$$\psi = \psi^{(0)} + \nu \psi^{(1)} + \nu^2 \psi^{(2)} + \cdots \tag{XXX-1}$$

$$E = E^{(0)} + \nu E^{(1)} + \nu^2 E^{(2)} + \cdots \tag{XXX-2}$$

The basic equations for determining these quantities are the familiar

$$(H^{(0)} - E^{(0)}) \psi^{(0)} = 0 \tag{XXX-3}$$

$$(H^{(0)} - E^{(0)}) \psi^{(1)} + (H^{(1)} - E^{(1)}) \psi^{(0)} = 0 \tag{XXX-4}$$

$$(H^{(0)} - E^{(0)}) \psi^{(2)} + (H^{(1)} - E^{(1)}) \psi^{(1)} + (H^{(2)} - E^{(2)}) \psi^{(0)} = 0 \tag{XXX-5}$$

etc.

As usual, we will assume that (3) has been solved for $\psi^{(0)}$ & $E^{(0)}$. Also we will assume that $\psi^{(0)}$ is normalized

$$(\psi^{(0)}, \psi^{(0)}) = 1 \qquad \text{(XXX-6)}$$

so that (4) yields

$$E^{(1)} = (\psi^{(0)}, H^{(1)} \psi^{(0)}) \qquad \text{(XXX-7)}$$

The next task then is to solve equation (4) for $\psi^{(1)}$. If we can do this then (5) yields

$$E^{(2)} = (\psi^{(0)}, H^{(1)} \psi^{(1)}) + (\psi^{(0)}, (H^{(0)} - E^{(0)}) \psi^{(2)}) \qquad \text{(XXX-8)}$$

and as we know, we can also calculate $E^{(3)}$ from $\psi^{(0)}$ and $\psi^{(1)}$.

Now for some problems one can solve (4) analytically or by a direct numerical approach[1]. However usually one must resort to further approximations. Following Hylleraas[2] we now observe that (4) can be derived from a variation principle

$$\delta J^{(2)}_H = 0 \quad , \text{all } \delta\psi^{(1)} \qquad \text{(XXX-9)}$$

where

$$J^{(2)}_H = (\psi^{(1)}, (H^{(0)} - E^{(0)}) \psi^{(1)}) + (\psi^{(1)}, (H^{(1)} - E^{(1)}) \psi^{(0)}) + (\psi^{(0)}, (H^{(1)} - E^{(1)}) \psi^{(1)}) \qquad \text{(XXX-10)}$$

and where, from (4) and (8), the numerical value of $J^{(2)}_H$ is evidently

$$J^{(2)}_H = E^{(2)} - (\psi^{(0)}, H^{(2)} \psi^{(0)}) \equiv E^{(2)\prime} \qquad \text{(XXX-11)}$$

These observations then immediately lead us to a variation method, the Hylleraas variation method, to approximate the solution of (4). Namely

$$\delta \hat{J}^{(2)}_{H} = 0 \qquad \text{all } \delta\hat{\psi}^{(1)} \tag{XXX-12}$$

where

$$\hat{J}^{(2)}_{H} = (\hat{\psi}^{(1)}, (H^{(0)} - E^{(0)})\hat{\psi}^{(1)}) + (\hat{\psi}^{(1)}, (H^{(1)} - E^{(1)})\psi^{(0)}) + (\psi^{(0)}, (H^{(1)} - E^{(1)})\hat{\psi}^{(1)}) \tag{XXX-13}$$

The numerical value of $\hat{J}^{(2)}_{H}$, which we will denote by $\hat{E}^{(2)'}$, then provides a variational approximation to $E^{(2)'}$, and with $\hat{\psi}^{(1)}$ we can also calculate a less well defined approximation to $E^{(2)}$. Written out in detail (12) becomes

$$(\delta\hat{\psi}^{(1)}, (H^{(0)} - E^{(0)})\hat{\psi}^{(1)} + (H^{(1)} - E^{(1)})\psi^{(0)}) + c.c. = 0 \tag{XXX-14}$$

We will discuss this variation method, and its generalizations in higher orders, in the next few sections. However first we would like to imbed it in a more general setting. Thus suppose we apply the variation principle

$$\delta E = 0 \qquad \text{all } \delta\psi \tag{XXX-15}$$

to trial functions of the form

$$\tilde{\psi} = \psi^{(0)} + \nu\,\tilde{\psi}^{(1)} + \nu^2\,\tilde{\psi}^{(2)} + \cdots$$

where the $\tilde{\psi}^{(n)}$ and their optimal values are assumed to be independent of ν but otherwise arbitrary. Since we have constrained their behavior for $\nu = 0$, such a set of trial functions is not the most general possible, but clearly it does contain ψ which is all we need.

Then

$$(\tilde{\psi}, (H-\tilde{E})\tilde{\psi}) = 0$$

yields

$$\tilde{E} = E^{(0)} + \nu E^{(1)} + \nu^2 \tilde{E}^{(2)} + \cdots \qquad \text{(XXX-16)}$$

where in particular

$$\tilde{E}^{(2)} = (\psi^{(0)}, H^{(2)}\psi^{(0)}) + \tilde{J}^{(2)}_H \qquad \text{(XXX-17)}$$

with

$$\tilde{J}^{(2)}_H = (\tilde{\psi}^{(1)}, (H^{(0)}-E^{(0)})\tilde{\psi}^{(1)}) + (\tilde{\psi}^{(1)}, (H^{(1)}-E^{(1)})\psi^{(0)}) + (\psi^{(0)}, (H^{(1)}-E^{(1)})\tilde{\psi}^{(1)}) \qquad \text{(XXX-18)}$$

and where, in deriving (16) and (18), we have made extensive use of (3) and (7). Eq. (11) is then an obvious consequence of (17) and since (15) implies

$$\delta E^{(2)} = 0 \qquad \text{also} \quad \delta\psi^{(1)} \qquad \text{(XXX-19)}$$

we also have rederived (9). [In accord with the notation which we introduced without comment in (9), we distinguish the optimal functions of the variation principle by simply deleting the $\sim$].

Similarly, starting with trial functions of the form[A]

$$\tilde{\psi} = \psi^{(0)} + \nu\psi^{(1)} + \nu^2\tilde{\psi}^{(2)} + \cdots \qquad \text{(XXX-20)}$$

one finds, in accord with the ν^{2n+1} theorem,

$$\tilde{E} = E^{(0)} + \nu E^{(1)} + \nu^2 E^{(2)} + \nu^3 E^{(3)} + \nu^4 \tilde{E}^{(4)} + \cdots \qquad \text{(XXX-21)}$$

and that

$$\delta E^{(4)} = 0 \qquad \text{all } \delta\psi^{(2)} \tag{XXX-22}$$

is equivalent to

$$\delta J^{(4)}_{H} = 0 \qquad \text{all } \delta\psi^{(2)} \tag{XXX-23}$$

where

$$\begin{aligned}\tilde{J}^{(4)}_{H} = {} & (\tilde{\psi}^{(2)}, (H^{(0)} - E^{(0)})\tilde{\psi}^{(2)}) + (\tilde{\psi}^{(2)}, (H^{(0)} - E^{(0)})\psi^{(0)}) \\ & + (\psi^{(0)}, (H^{(1)} - E^{(1)})\tilde{\psi}^{(2)}) + (\tilde{\psi}^{(2)}, (H^{(2)} - E^{(2)})\psi^{(0)}) \\ & + (\psi^{(0)}, (H^{(2)} - E^{(2)})\tilde{\psi}^{(2)})\end{aligned} \tag{XXX-24}$$

and where the numerical value of $J^{(4)}_{H}$ is

$$\begin{aligned}J^{(4)}_{H} = {} & E^{(4)} - (\psi^{(0)}, H^{(4)}\psi^{(0)}) - (\psi^{(0)}, (H^{(3)} - E^{(3)})\psi^{(1)}) \\ & - (\psi^{(1)}, (H^{(3)} - E^{(3)})\psi^{(0)}) \equiv E^{(4)\prime}\end{aligned} \tag{XXX-25}$$

The variation principle (23) is of course equivalent to equation (5) and could have been written down directly just as we wrote down (9). However, again, it is nice to see it as part of a larger setting.

Given $\psi^{(0)}$ and $\psi^{(0)}$ then, the equations

$$\delta \hat{J}^{(4)}_{H} = 0 \qquad \text{"all" } \delta\hat{\psi}^{(2)} \tag{XXX-26}$$

provide a variation method for approximating the solution (5), with the numerical value of $\hat{J}^{(4)}_{H}$ providing a variational approximation to $E^{(4)\prime}$. Similarly one can go on to $\hat{J}^{(6)}_{H}$ etc.; however since the procedure is obvious, we will not write out the details.

In Sec. XXVIII, following Eq. (32), we made a rather cryptic comment which we will now attempt to explain. Suppose that we are doing perturbation theory within the linear variation method. Then since this is Rayleigh-Schroedinger perturbation theory for $\overline{H}$ we can apply the results of the present section. Thus the basic equations for $\hat{a}_k^{(0)}$, $\hat{a}_k^{(1)}$ etc. can be derived from

$$\delta \overline{J}_{kH}^{(2)} = 0 \qquad \text{all } \delta\hat{\psi}_k^{(1)} \tag{XXX-27}$$

$$\delta \overline{J}_{kH}^{(4)} = 0 \qquad \text{all } \delta\hat{\psi}_k^{(2)} \tag{XXX-28}$$

etc., where

$$\overline{J}_{kH}^{(2)} = (\hat{\psi}_k^{(1)}, (\overline{H}^{(0)} - \hat{E}_k^{(0)})\hat{\psi}_k^{(1)}) + (\hat{\psi}_k^{(1)}, (\overline{H}^{(1)} - \hat{E}_k^{(1)})\hat{\psi}_k^{(0)}) + (\hat{\psi}_k^{(0)}, (\overline{H}^{(1)} - \hat{E}_k^{(1)})\hat{\psi}_k^{(1)}) \tag{XXX-29}$$

$$\overline{J}_{kH}^{(4)} = (\hat{\psi}_k^{(3)}, (\overline{H}^{(0)} - \hat{E}_k^{(0)})\hat{\psi}_k^{(3)}) + (\hat{\psi}_k^{(3)}, (\overline{H}^{(0)} - \hat{E}_k^{(0)})\hat{\psi}_k^{(1)}) + (\hat{\psi}_k^{(1)}, (\overline{H}^{(0)} - \hat{E}_k^{(0)})\hat{\psi}_k^{(3)}) + (\hat{\psi}_k^{(0)}, (\overline{H}^{(1)} - \hat{E}_k^{(1)})\hat{\psi}_k^{(2)}) + (\hat{\psi}_k^{(2)}, (\overline{H}^{(1)} - \hat{E}_k^{(1)})\hat{\psi}_k^{(0)}) + (\hat{\psi}_k^{(0)}, (\overline{H}^{(2)} - \hat{E}_k^{(2)})\hat{\psi}_k^{(0)}) \tag{XXX-30}$$

etc.

Possible confusion now arises in the common case that the set of trial functions is independent of ν and that (recall Eq. (XXVIII-1) and (XXVIII-27)

$$\overline{H}^{(0)} = \overline{H^{(0)}} \quad , \quad \overline{H}^{(1)} = \overline{H^{(1)}} \quad , \quad \overline{H}^{(2)} = \overline{H^{(2)}} \quad \text{etc.} \tag{XXX-31}$$

Then since the $\hat{\psi}^{(n)}_k$ are all within the linear space we may replace the $\overline{H}^{(n)}$ by $H^{(n)}$ in (29), (30) etc. and hence instead of (27)-(30) we can write

$$\delta \overline{J}^{(n)}_{kn} = 0 \qquad \text{all } \delta\hat{\psi}^{(n)} \text{ in the space} \tag{XXX-32}$$

where

$$\overline{J}^{(n)}_{kn} = (\hat{\psi}^{(n)}_k, (H^{(0)} - \hat{E}^{(0)}_k)\hat{\psi}^{(n)}_k) + (\hat{\psi}^{(n)}_k, (H^{(0)} - \hat{E}^{(0)}_k)\hat{\psi}^{(0)}_k) + (\hat{\psi}^{(0)}_k, (H^{(0)} - \hat{E}^{(0)}_k)\hat{\psi}^{(n)}_k) \tag{XXX-33}$$

and

$$\delta \overline{J}^{(n)}_{kn} = 0 \qquad \text{all } \delta\hat{\psi}^{(n)} \text{ in the space} \tag{XXX-34}$$

where

$$\overline{J}^{(n)}_{kn} = (\hat{\psi}^{(n)}_k, (H^{(0)} - \hat{E}^{(0)}_k)\hat{\psi}^{(n)}_k) + (\hat{\psi}^{(n)}_k, (H^{(0)} - \hat{E}^{(0)}_k)\hat{\psi}^{(0)}_k) + (\hat{\psi}^{(0)}_k, (H^{(0)} - \hat{E}^{(0)}_k)\hat{\psi}^{(n)}_k) + (\hat{\psi}^{(n)}_k, (H^{(n)} - \hat{E}^{(n)}_k)\hat{\psi}^{(0)}_k) + (\hat{\psi}^{(0)}_k, (H^{(n)} - \hat{E}^{(n)}_k)\hat{\psi}^{(n)}_k) \tag{XXX-35}$$

The point is now that making due allowance for the different notation, comparison of (32)-(35) with (12), (13), and (26), (24) might then suggest that successive terms in the expansion of $\hat{E}_k$ yield variational approximations to the $E^{(n)}$ of some state k. However it should be obvious that (32) and (33) are a special case of (12) and (13) only if $\hat{\psi}^{(0)}_k$ is a $\psi^{(0)}$, i.e., an eigenfunction of $H^{(0)}$; and that (34) and (35) are a special case of (26) and (24) only if in addition $\hat{\psi}^{(1)}_k$ is the corresponding $\psi^{(1)}$. Thus in general the $\hat{E}^{(n)}_k$ and $\hat{\psi}^{(n)}_k$ are simply the successive terms in the expansion of E_k within the linear variation method and nothing

more. In particular, unless special conditions are satisfied, they have no particular relationship to the $E^{(n)}$ and $\psi^{(n)}$.

XXXI. THE HYLLERAAS VARIATION METHOD - SOME GENERAL RESULTS

Since $\delta J^{(1)}_H = \delta J^{(2)}_H = \cdots = 0$ it follows that if we write

$$\hat{\psi}^{(n)} = \psi^{(n)} + \Delta^{(n)}$$

then the terms of first order in $\Delta^{(n)}$ must cancel out when we compute $\hat{J}^{(2n)}_H - J^{(2n)}_H$ so that we will be left with

$$\hat{J}^{(2n)}_H - J^{(2n)}_H = (\Delta^{(n)}, (H^{(0)} - E^{(0)})\Delta^{(n)}) \qquad \text{(XXXI-1)}$$

From this result several useful conclusions can be drawn:

(i) If $\psi^{(0)}$ and ψ share a common symmetry, for example both satisfy the Pauli Principle, and if $\psi^{(0)}$ is the lowest state of $H^{(0)}$ with that symmetry then if the $\hat{\psi}^{(n)}$, and hence $\Delta^{(n)}$, also have that symmetry it follows from (1) that

$$\hat{J}^{(2n)}_H \geq J^{(2n)}_H \qquad \text{(XXXI-2)}$$

That is $\hat{J}^{(2n)}_H$ will provide an upper bound to $E^{(2n)}$. In such cases then the Hylleraas variatiom method is very well founded, and capable of systematic improvement.

(ii) In many problems $\psi^{(0)}$ will have some symmetry property not shared by ψ . If then in a given order, all the eigenstates of $H^{(0)}$ with appropriate symmetry to contribute to $\psi^{(n)}$ have energies greater than $E^{(0)}$, we will again have (2) if the $\hat{\psi}^{(n)}$ have the appropriate symmetry.

In either case (i) or (ii) we will refer to $\psi^{(0)}$ as the "ground state", though the name is rather more appropriate to case (i).

A second proof of (2) in case (i) (the proof can also be adapted to case (ii) - see Sec. XXXIV) follows from the observation that if ψ is the ground state for $V=0$ it will also be the ground state for V sufficiently small. Thus if we use a trial function

$$\tilde{\psi} = \psi^{(0)} + V\psi^{(1)} + V^2\psi^{(2)} + \cdots \tag{XXXI-3}$$

in the variation method, then for V small enough we must find

$$\tilde{E} \geq E \tag{XXXI-4}$$

Expansion in powers of V and use of the V^{2n+1} theorem then yields

$$E^{(0)} + V E^{(1)} + V^2 E^{(2)} + \cdots \geq E^{(0)} + V E^{(1)} + V^2 E^{(2)} + \cdots \tag{XXXI-5}$$

and therefore

$$\tilde{E}^{(n)} \geq E^{(n)} \tag{XXXI-6}$$

which, in turn yields

$$\tilde{E}^{(n)\prime} = \tilde{J}^{(n)}_{11} \geq E^{(n)\prime} = J^{(n)}_{11} \tag{XXXI-7}$$

Similarly

$$\tilde{\psi} = \psi^{(0)} + V\psi^{(1)} + V^2\psi^{(2)} + \cdots \tag{XXXI-8}$$

yields

$$E^{(0)} + V E^{(1)} + V^2 E^{(2)} + V^3 E^{(3)} + V^4 E^{(4)} + \cdots \geq E^{(0)} + V E^{(1)} + V^2 E^{(2)} + V^3 E^{(3)} + V^4 E^{(4)} + \cdots \tag{XXXI-9}$$

and hence

$$\hat{E}^{(4)} \geq E^{(4)} \qquad \text{(XXXI-10)}$$

etc. We will discuss procedures which yield upper bounds on $E^{(2)}$, $E^{(4)}$ etc. for excited states in Sec. XXXIV.

The numerical value of $J_4^{(2)}$ can be written in various equivalent forms. Thus starting from (10) and using (XXX-4) one readily derives the alternative forms (the first having been given already in (XXX-8))

$$E^{(2)'} = (\psi^{(0)}, (H^{(0)} - E^{(0)})\psi^{(0)}) \qquad \text{(XXXI-11)}$$

$$E^{(2)'} = (\psi^{(1)}, (H^{(0)} - E^{(0)})\psi^{(0)}) \qquad \text{(XXXI-12)}$$

$$E^{(2)'} = -(\psi^{(1)}, (H^{(0)} - E^{(0)})\psi^{(1)}) \qquad \text{(XXXI-13)}$$

The form (12) can of course be derived from (11) and the fact that $E^{(2)'}$ is real. The form (13) makes it obvious that $E^{(2)'}$ is ≤ 0 for the ground state.

Similar alternative formulae can be derived for $E^{(2)'}$ if the $\hat{\psi}^{(1)}$ have no fixed overall size. Namely this will imply that

$$\delta a\, \hat{\psi}^{(0)}$$

with δa an infinitesimal but otherwise arbitrary complex number will be a possible variation of $\hat{\psi}^{(0)}$ within the set. Eq. (XXX-14) with δa successively real and pure imaginary then yields

$$2(\hat{\psi}^{(1)}, (H^{(0)} - E^{(0)})\hat{\psi}^{(1)}) + (\hat{\psi}^{(1)}, (H^{(0)} - E^{(0)})\psi^{(0)}) + (\psi^{(0)}, (H^{(0)} - E^{(0)})\hat{\psi}^{(1)}) = 0 \qquad \text{(XXXI-14)}$$

and

$$(\hat{\psi}^{(1)}, (H^{(0)} - E^{(0)})\psi^{(0)}) - (\psi^{(0)}, (H^{(0)} - E^{(0)})\hat{\psi}^{(1)}) = 0 \qquad \text{(XXXI-15)}$$

These results combined with (XXX-13) then yield the alternative forms

$$\hat{E}^{(2)\prime} = (\psi^{(0)}, (H^{(1)} - E^{(1)})\hat{\psi}^{(0)}) \qquad \text{(XXXI-16)}$$

$$\hat{E}^{(2)\prime} = (\hat{\psi}^{(1)}, (H^{(1)} - E^{(1)})\psi^{(0)}) \qquad \text{(XXXI-17)}$$

$$\hat{E}^{(2)\prime} = -(\hat{\psi}^{(1)}, (H^{(0)} - E^{(0)})\hat{\psi}^{(1)}). \qquad \text{(XXXI-18)}$$

In particular the form (18) shows that under these conditions $\hat{E}^{(2)\prime}$ will be ≤ 0 for the ground state. With similar assumptions about the $\hat{\psi}^{(2)}$, etc. analogous rewritings of $\hat{E}^{(2n)\prime}$ etc. can be made; however we will not pursue the details.

When the set of trial functions $\hat{\psi}^{(0)}$ has no fixed overall scale then it is possible to rewrite the basic equations (12) and (13) in an interesting and sometimes useful way (again analogous results hold for $\hat{J}^{(H)}_{H}$ etc.). Namely let us write

$$\hat{\psi}^{(1)} = \hat{N}\,\hat{\phi}^{(1)} \qquad \text{(XXXI-19)}$$

where $\hat{N}$ is a constant. Then since

$$\delta \hat{J}^{(2)}_{H} = 0 \qquad \text{all} \quad \delta\hat{N}$$

Eq. (XXX-14) yields

$$(\delta\hat{N})^{*}\hat{N}(\hat{\phi}^{(1)}, (H^{(0)} - E^{(0)})\hat{\phi}^{(1)}) + (\delta\hat{N})^{*}(\hat{\phi}^{(1)}, (H^{(1)} - E^{(1)})\psi^{(0)}) = 0 \qquad \text{(XXXI-20)}$$

and therefore

$$\hat{N} = -\frac{(\hat{\phi}^{(1)}, (H^{(1)} - E^{(1)})\psi^{(0)})}{(\hat{\phi}^{(1)}, (H^{(0)} - E^{(0)})\hat{\phi}^{(1)})} \qquad \text{(XXXI-21)}$$

If we now use this value of $\hat{N}$ in the Hylleraas functional $\hat{J}^{(2)}_H$ we get a new functional $\hat{J}^{(2)}_D$, the D standing for Dirichlet

$$\hat{J}^{(2)}_D = - \frac{|(\hat{\Phi}^{(1)}, (H^{(1)} - E^{(1)})\Psi^{(0)})|^2}{(\hat{\Phi}^{(1)}, (H^{(0)} - E^{(0)})\hat{\Phi}^{(1)})} \qquad \text{(XXXI-22)}$$

and we can replace (12) by

$$\delta \hat{J}^{(2)}_{D=0} \quad \text{"all"} \; \delta\hat{\Phi}^{(1)} \qquad \text{(XXXI-23)}$$

Although we have derived (23) from (12)[A] there is a certain sense in which the formulation (23) is "better".[1] Namely (23) is independent of the scale of the $\hat{\Phi}^{(1)}$. Therefore if we would draw the $\tilde{\Psi}^{(1)}$ and the $\tilde{\Phi}^{(1)}$ from the same set of <u>fixed</u> scale trial functions (23) will in general yield better results (a better upper bound to $E^{(2)}$, in the case of the ground state) since these results are what one would get using $\hat{J}^{(2)}_H$ if one <u>enlarged</u> the set of $\tilde{\Psi}^{(1)}$ by allowing their overall scale to be arbitrary.

Throughout our discussion we have assumed that $\Psi^{(0)}$ is known. However often this is not the case. Thus if we are interested in calculating the polarizability of an atom or molecule then for $H^{(0)}$ we should take the Hamiltonian of the isolated system and $H^{(1)}$ should be the interaction with the external field. However only for very simple systems do we then know $\Psi^{(0)}$ with great accuracy. Instead one usually has some (normalized) approximation $\chi^{(0)}$, often of dubious quality.

There have been several responses made to this situation. One is to view $\chi^{(0)}$ as defining a "model system". Thus one everywhere

replaces $H^{(0)}$ by a "Model" Hamiltonian $H^{(00)}$ of which $\chi^{(0)}$ is an eigenfunction and proceeds with the analysis;[B] the difference $H^{(0)}-H^{(00)}$ perhaps to be used subsequently as a further perturbation in hopes of improving the results. (See Sec. XXXVI).

Another approach which has been suggested[2] in case $\psi^{(0)}$ is the ground state, is to try to replace the numerator in (22) by some known upper bound; so that the result is still an upper bound to $E^{(2)}$. Details may be found in the reference cited. A third approach, and one which we will now discuss in more detail, is simply to apply (XXX-12) and (XXX-13) but with $\psi^{(0)}$ replaced by $\chi^{(0)}$. (Similarly one might try to use $\hat{T}^{(4)}_H$, $\hat{T}^{(6)}_H$ etc. but with $\psi^{(0)}$ and $\psi^{(1)}$ etc. replaced by approximations). Thus one requires

$$\delta \dot{J} = 0 \qquad \text{"all" } \delta\dot{\psi}^{(1)} \tag{XXXI-24}$$

where

$$\dot{J} = (\dot{\psi}^{(1)}, (H^{(0)}-\epsilon^{(0)})\dot{\psi}^{(1)}) + (\dot{\psi}^{(1)}, (H^{(1)}-\epsilon^{(1)})\chi^{(0)}) + (\chi^{(0)}, (H^{(1)}-\epsilon^{(1)})\dot{\psi}^{(1)}) \tag{XXXI-25}$$

and

$$\epsilon^{(0)} = (\chi^{(0)}, H^{(0)}\chi^{(0)}) \quad , \quad \epsilon^{(1)} = (\chi^{(0)}, H^{(1)}\chi^{(0)}) \tag{XXXI-26}$$

$$(\chi^{(0)}, \chi^{(0)}) = 1 \tag{XXXI-27}$$

The functional $\dot{J}$ is of course much like the functional which appears in (XXVI-45), indeed we are here concerned with the same problem which we discussed there. However we have evidently by-passed anything like Eq. (XXVI-48) and also there is no analogue of $\tilde{\chi}^{(2)}$ in (25). (Such

terms would of course play no role in the variation if $(H^{(0)} - \epsilon^{(0)})\chi^{(0)}$ were equal to zero).[3] We first consider situations in which it is known, usually for reasons of symmetry that

$$(\psi^{(0)}, \dot{\psi}^{(1)}) = 0 \qquad \text{(XXXI-28)}$$

An analysis of (25) and (26), which we will not reproduce in detail here,[4] then yields the following results,

(i) Writing $\chi^{(0)} = \psi^{(0)} + \Delta^{(0)}$ and $\dot{\psi}^{(1)} = \psi^{(1)} + \Delta^{(1)}$ one finds that the leading term (at least in a formal sense) in $\dot{J} - \epsilon^{(2)}$ is of order $\Delta^{(0)}$ and may have either sign. However one can hope that $\chi^{(0)}$ is sufficiently accurate that its numerical value is small.

(ii) If $\epsilon^{(0)}$ lies between the ground state eigenvalue of $H^{(0)}$ and the eigenvalue of the lowest eigenstate of $H^{(0)}$ which is not orthogonal to the $\chi^{(0)}$ then $\dot{J}$ is an upper bound on its "exact" value - the value which we obtain when $\dot{\psi}^{(1)}$ satisfies the equation derived from $\delta \dot{J} = 0$ all $\delta\dot{\psi}^{(1)}$, i.e.,

$$(H^{(0)} - \epsilon^{(0)})\dot{\psi}^{(1)} + (H^{(1)} - \epsilon^{(1)})\chi^{(0)} = 0 \qquad \text{(XXXI-29)}$$

Thus, in this case, the value of $\dot{J}$ will decrease (or at least not increase) as one makes the set of trial functions more flexible.

(iii) In many practical cases $\epsilon^{(0)}$, far from being low lying in the discrete spectrum of $H^{(0)}$, may be in the continuum. In such situations (24) will lead to nonsense if one makes the set of trial functions too flexible, with undefined integrals and the like appearing. That the procedure nevertheless often yields reasonable results in

practice seems to be due to a cancellation of errors, the smallness of $\chi^{(0)}$ being balanced by the relative inflexibility of the set of trial functions $\varphi^{(0)}$.

Turning now to the case in which $(\psi^{(0)}, \dot{\psi}^{(1)})$ is not automatically zero by symmetry, the major point to be made is that (29) should be replaced by

$$\dot{J} = (\dot{\psi}^{(1)}, (H^{(0)} - \epsilon^{(0)})\dot{\psi}^{(1)}) + (\dot{\psi}^{(1)}, (H^{(1)} - \dot{\xi})\chi^{(0)}) + (\chi^{(0)}, (H^{(1)} - \dot{\xi})\dot{\psi}^{(1)}) \qquad \text{(XXXI-30)}$$

where the Lagrange multiplier $\dot{\xi}$ is to be chosen so that

$$(\chi^{(0)}, \dot{\psi}^{(1)}) = 0 \qquad \text{(XXXI-31)}$$

The reason for this is that if we choose the "best possible" $\dot{\psi}^{(1)}$ in the sense of (XXXI-24) then we are led to (29), and hence to

$$(E^{(0)} - \epsilon^{(0)})(\psi^{(0)}, \dot{\psi}^{(1)}) + (\psi^{(0)}, (H^{(1)} - \mathcal{E}^{(1)})\chi^{(0)}) = 0 \qquad \text{(XXXI-32)}$$

However as

$$\chi^{(0)} - \psi^{(0)} \equiv \Delta^{(0)} \qquad \text{(XXXI-33)}$$

tends to zero ($E^{(0)} - \epsilon^{(0)}$) is of order $(\Delta^{(0)})^2$ while naturally ($\psi^{(0)}$, $(H^{(1)} - \mathcal{E}^{(1)})\chi^{(0)}$) is of order $\Delta^{(0)}$ so that

$$(\psi^{(0)}, \dot{\psi}^{(1)}) \sim \sqrt{\Delta^{(0)}} \qquad \text{(XXXI-34)}$$

Thus the solution of (29) becomes ill defined just in the limit $\Delta^{(0)} \to 0$ when it should be becoming exact. On the other hand if we replace $\mathcal{E}^{(1)}$ by $\dot{\xi}$, the latter to be chosen so that (32) is satisfied, then there are no difficulties in the limit $\Delta^{(0)} \to 0$.

(iv) With the definition (29) then (i) holds also in the present case (if one uses (25) there is an error of order zero!).

(v) If $\epsilon^{(0)}$ is less than or equal to the lowest eigenvalue of $\Pi H^{(0)} \Pi + (1-\Pi) H^{(0)} (1-\Pi)$ where Π is the projector onto $\chi^{(0)}$, then $\dot{J}$ will decrease, (or at any rate not increase) as one makes the set of $\tilde{\psi}^{(1)}$ more flexible.[D]

XXXII. THE LINEAR HYLLERAAS VARIATION METHOD

Trial functions of the form

$$\tilde{\psi}^{(n)} = \sum_{1}^{M} \tilde{a}_{\kappa}^{(n)} \Phi_{\kappa} \qquad \text{(XXXII-1)}$$

with the Φ_{κ} a given basis set (which may contain non-linear parameters; however we will confine our attention to the linear parameters) are often used in the Hylleraas variation method, and in this section we will write down some of the basic equations of this approach. We will confine our attention to $n=1$, however the results obviously generalize. For this purpose we could, to a large extent, take over the results of Sec. XXVIII specialized to the case in which $\hat{\psi}^{(0)} = \psi^{(0)}$. However, though it is somewhat repetitious, let us start afresh.

Thus with

$$\hat{\psi}^{(1)} = \sum_{\kappa=1}^{M} \hat{a}_{\kappa}^{(1)} \Phi_{\kappa} \qquad \text{(XXXII-2)}$$

$$\delta\hat{\psi}^{(1)} = \sum_{\kappa=1}^{M} \delta\hat{a}_{\kappa}^{(1)} \Phi_{\kappa} \qquad \text{(XXXII-3)}$$

equation (XXX-14) yields

$$\sum_L (\Phi_k, (H^{(0)} - E^{(0)})\Phi_L)\hat{C}_L^{(1)} + (\Phi_k, (H^{(1)} - E^{(1)})\psi^{(0)}) = 0 \qquad \text{(XXXII-4)}$$

This is essentially equation (XXVIII-II) for the case $S^{(1)} = 0$ with, however, the difference that $\psi^{(0)}$ itself need not be expandable in the Φ_k, i.e., $\psi^{(0)}$ need not be a possible $\hat{\psi}^{(1)}$.

For the case in which $\psi^{(0)}$ is a possible $\hat{\psi}^{(1)}$ we can simply take over the results of Sec. XXVIII and hence, in particular, from the discussion following (XXVIII-2) we can infer that

$$\hat{\psi}^{(1)} = -\sum_L \hat{\psi}_L^{(0)} \frac{(\hat{\psi}_L^{(0)}, H^{(1)}\psi^{(0)})}{\hat{E}_L^{(0)} - E^{(0)}} \qquad \text{(XXXII-5)}$$

and

$$\hat{E}^{(1)} = -\sum_L \frac{|(\hat{\psi}_L^{(0)}, H^{(1)}\psi^{(0)})|^2}{\hat{E}_L^{(0)} - E^{(0)}} \qquad \text{(XXXII-6)}$$

where the (M-1) functions $\hat{\psi}_L^{(0)}$ together with $\psi^{(0)}$ are the orthonormal eigenfunctions of $H^{(0)}$ within the basis set, the $\hat{E}_L^{(0)}$ and $E^{(0)}$ then being the corresponding eigenvalues. Thus the sums in (5) and (6) each contains at most (M-1) terms. Also in writing down (5) and (6) we have assumed that none of the $\hat{E}_L^{(0)}$ equals $E^{(0)}$. (If $E^{(0)}$ is non-degenerate then clearly such an accidental degeneracy can occur only if $\psi^{(0)}$ is an excited state. Also if one does have such an accidental degeneracy then the equations (4) become inconsistent unless the corresponding $(\hat{\psi}_L^{(0)}, H^{(1)}\psi^{(0)})$ happens to vanish).

Turning now to the case in which $\psi^{(0)}$ is not a possible $\hat{\psi}^{(1)}$ we first observe that if we were to use the Φ_k in a linear variation calculation for $H^{(0)}$ we would be led to the equation

$$\left| \left(\phi_k, \left(H^{(0)} - \dot{E}^{(0)} \right) \phi_L \right) \right| = 0 \qquad \text{(XXXII-7)}$$

to determine the optimal energies $\dot{E}^{(0)}_k$. Then if, as we will assume, none of the $\dot{E}^{(0)}_k$ equals $E^{(0)}$ (and if $\psi^{(0)}$ is the ground state its absence from among the $\tilde{\psi}^{(k)}$ is already sufficient to ensure this) it will follow that the $M \times M$ matrix $\mathfrak{M}$ whose matrix elements are

$$\mathfrak{M}_{kL} = \left(\phi_k, \left(H^{(0)} - E^{(0)} \right) \phi_L \right) \qquad \text{(XXXII-8)}$$

will have an inverse. Therefore from (4) we will have

$$\hat{a}^{(1)}_k = - \sum_L \left(\mathfrak{M}^{-1} \right)_{kL} \left(\phi_L, \left(H^{(1)} - E^{(1)} \right) \psi^{(0)} \right) \qquad \text{(XXXII-9)}$$

and hence, from (XXXI-17)

$$\hat{E}^{(2)\prime} = - \sum_k \sum_L \left(\psi^{(0)}, \left(H^{(1)} - E^{(1)} \right) \phi_k \right) \left(\mathfrak{M}^{-1} \right)_{kL} \left(\phi_L, \left(H^{(1)} - E^{(1)} \right) \psi^{(0)} \right) \qquad \text{(XXXII-10)}$$

For M = 1, Eq. (10) is, of course, just the Dirichlet functional of (XXXI-22)^A.

To compare with (5) and (6) it is useful to introduce the M functions which are the eigenfunctions of $\overline{H^{(0)}}$ within the basis set. If we denote these functions by $\dot{\psi}^{(0)}_k$ and use them as the ϕ_k, then $\mathfrak{M}$ is a diagonal matrix and (9) and (10) yield

$$\hat{\psi}^{(1)} = - \sum_{L=1}^{M} \dot{\psi}^{(0)}_L \frac{\left(\dot{\psi}^{(0)}_L, \left(H^{(1)} - E^{(1)} \right) \psi^{(0)} \right)}{\dot{E}^{(0)}_L - E^{(0)}} \qquad \text{(XXXII-11)}$$

and

$$\hat{E}^{(2)\prime} = - \sum_{L=1}^{M} \frac{\left| \left(\psi^{(0)}, \left(H^{(1)} - E^{(1)} \right) \dot{\psi}^{(0)}_L \right) \right|^2}{\dot{E}^{(0)}_L - E^{(0)}} \qquad \text{(XXXII-12)}$$

respectively. These formulae are quite similar to (5) and (6) except that the sums in general contain M terms.

The need for slightly different procedures in the two cases is of course basically due to the fact that $(H^{(0)} - E^{(0)})$ has a zero eigenvalue and therefore has no operator inverse. Formally we could have avoided this difficulty by replacing $E^{(0)}$ by $E^{(0)} - \epsilon$ where ϵ is a small number. With this replacement, and continuing to assume no degeneracy we can simply apply (11) and (12) to the general case. If then $\psi^{(0)}$ is not one of the $\psi_L^{(0)}$ then when we let $\epsilon \to 0$ we will regain (11) and (12). On the other hand if $\psi^{(0)}$ is one of the $\psi_L^{(0)}$ its contribution to (11) and (12) will vanish identically since

$$\frac{(\psi^{(0)}, (H^{(1)} - E^{(1)})\psi^{(0)})}{\epsilon} = 0 \qquad \epsilon \neq 0 \tag{XXXII-13}$$

and hence if we then let $\epsilon \to 0$ we will recover (5) and (6).

If ϵ is in addition a positive number and if $\psi^{(0)}$ is the ground state then even for $\epsilon \neq 0$ the quantity

$$\tilde{J}_\epsilon \equiv (\tilde{\psi}^{(1)}, (H^{(0)} - E^{(0)} + \epsilon)\tilde{\psi}^{(1)}) + (\tilde{\psi}^{(1)}, (H^{(1)} - E^{(1)})\psi^{(0)}) + (\psi^{(0)}, (H^{(1)} - E^{(1)})\tilde{\psi}^{(1)})$$

is evidently an upper bound on its "exact" value because $(H^{(0)} - E^{(0)} + \epsilon)$ is a strictly positive operator. By "exact value" here we mean its numerical value if one uses for $\tilde{\psi}^{(1)}$ the function $\zeta^{(1)}$ which satisfies

$$(H^{(0)} - E^{(0)} + \epsilon)\zeta^{(1)} + (H^{(1)} - E^{(1)})\psi^{(0)} = 0$$

More generally consider the inhomogeneous equation

$$-\Omega\zeta + g = 0 \tag{XXXII-14}$$

where Ω is a strictly positive operator and where g is some known function. Then, by the standard arguments

$$\hat{J} = (\hat{\xi}, \Omega \hat{\xi}) + (\hat{\xi}, g) + (g, \hat{\xi}) \qquad \text{(XXXII-15)}$$

will be an upper bound to

$$(g, \xi) = -(g, \Omega^{-1} g) \qquad \text{(XXXII-16)}$$

A linear variation calculation with

$$\hat{\xi} = \sum_{k=1}^{M} c_k \phi_k$$

then yields

$$\hat{J} = -\sum_k \sum_l (g, \phi_l)(\mathcal{A}^{-1})_{lk}(\phi_k, g) \qquad \text{(XXXII-17)}$$

where $\mathcal{A}$ is an $M \times M$ matrix whose matrix elements are

$$\mathcal{A}_{kl} = (\phi_k, \Omega \phi_l) \qquad \text{(XXXII-18)}$$

Moreover since the result

$$\hat{J} \geq -(g, \Omega^{-1} g)$$

is true for any g then we have the interesting operator inequality[1]

$$\Omega^{-1} \geq \sum_k \sum_l |\phi_l\rangle (\mathcal{A}^{-1})_{lk} \langle \phi_k| \qquad \text{(XXXII-19)}$$

In Appendix C we will derive a generalization of this result which holds when Ω has a finite number of negative eigenvalues.[B]

The fact that $(H^{(0)} - E^{(0)})$ has a zero eigenvalue, i.e., the fact that

$$(H^{(0)} - E^{(0)})\psi^{(0)} = 0 \qquad \text{(XXXII-20)}$$

also has another interesting and possibly confusing consequence. Suppose that we start with a set of B basis functions which does not include $\psi^{(0)}$. Then the preceding analysis will lead us to (11) and (12) with $M = B$. However now we note that if we add $\psi^{(0)}$ to the set we must get exactly the same results since clearly from the form of $\tilde{F}^{(2)}_H$, $\tilde{\psi}^{(1)}$ and $\tilde{\psi}^{(1)} + \alpha\psi^{(0)}$ yield exactly the same numerical values for $\tilde{F}^{(2)}_H$. However with $\psi^{(0)}$ in the set the preceding analysis would lead us to (5) and (6) though now with $M = B+1$ and the point is that at this stage, it is no longer obvious that the results agree.

Although it is not difficult to give a detailed proof that the two sets of formulae in fact do agree, we will confine attention to the case $B=1$ which suffices to bring out the main points. Denoting the single normalized basis function by ϕ equation (11) and (12) become

$$\tilde{\psi}^{(1)} = -\phi \frac{(\phi, (H^{(1)} - E^{(1)})\psi^{(0)})}{(\phi, H^{(0)}\phi) - E^{(0)}} \qquad \text{(XXXII-21)}$$

and

$$\tilde{E}^{(2)} = -\frac{|(\phi, (H^{(1)} - E^{(1)})\psi^{(0)})|^2}{(\phi, H^{(0)}\phi) - E^{(0)}} \qquad \text{(XXXII-22)}$$

respectively.

On the other hand it is clear that the single $\tilde{\psi}^{(0)}_1$ of Eq. (5) and (6), call it $\tilde{\psi}^{(0)}_1$, is

$$\tilde{\psi}^{(0)}_1 = \frac{\phi - \psi^{(0)}(\psi^{(0)}, \phi)}{\sqrt{1 + |(\psi^{(0)}, \phi)|^2}} \qquad \text{(XXXII-23)}$$

since this function is normalized, and is orthogonal to $\psi^{(0)}$. Then since

$$\hat{E}_1^{(0)} = (\hat{\psi}_1^{(0)}, H^{(0)} \hat{\psi}_1^{(0)}) \qquad \text{(XXXII-24)}$$

we find

$$\hat{E}_1^{(0)} = \frac{(\phi, H^{(0)}\phi) + E^{(0)} |(\psi^{(0)}, \phi)|^2}{1 + |(\psi^{(0)}, \phi)|^2} \qquad \text{(XXXII-25)}$$

and also we have

$$(\hat{\psi}_1^{(0)}, H^{(1)} \psi^{(0)}) = \frac{(\phi, (H^{(1)} - E^{(1)}) \psi^{(0)})}{\sqrt{1 + |(\psi^{(0)}, \phi)|^2}} \qquad \text{(XXXII-26)}$$

With these results (6) then yields

$$E^{(2)} = - \frac{|(\phi, (H^{(1)} - E^{(1)}) \psi^{(0)})|^2}{1 + |(\psi^{(0)}, \phi)|^2} \Big/ \left(\frac{(\phi, H^{(0)}\phi) + E^{(0)} |(\psi^{(0)}, \phi)|^2}{1 + |(\psi^{(0)}, \phi)|^2} - E^{(0)} \right) \qquad \text{(XXXII-27)}$$

which is readily seen to be identical to (22) as it must be. Similarly one finds that (5) yields a $\hat{\psi}^{(1)}$ which is essentially identical to (21), i.e., they differ only by a physically irrelevant multiple of $\psi^{(0)}$.

XXXIII. THE SECOND ORDER HYLLERAAS VARIATION METHOD WITH TWO PERTURBATIONS[1]

Suppose that there are two first-order perturbations. Then with the replacements

$$\cdot \; H^{(1)} \longrightarrow V_1 H_1^{(1)} + V_2 H_2^{(1)} \qquad \text{(XXXIII-1)}$$

$$V^2 H^{(2)} \rightarrow v_1^2 H_{11}^{(2)} + 2v_1 v_2 H_{12}^{(2)} + v_2^2 H_{22}^{(2)} \qquad \text{(XXXIII-2)}$$

$$V E^{(1)} \rightarrow v_1 E_1^{(1)} + v_2 E_2^{(1)} \qquad \text{(XXXIII-3)}$$

$$V^2 E^{(2)} \rightarrow v_1^2 E_{11}^{(2)} + 2v_1 v_2 E_{12}^{(2)} + v_2^2 E_{22}^{(2)} \qquad \text{(XXXIII-4)}$$

$$V \psi^{(1)} \rightarrow v_1 \psi_1^{(1)} + v_2 \psi_2^{(1)} \qquad \text{(XXXIII-5)}$$

The latter three being valid only if there are no problems with degeneracy, one readily finds from the appropriate equations of Sec. XXX that the $\psi_i^{(1)}$ are determined by

$$(H^{(0)} - E^{(0)}) \psi_i^{(1)} + (H_i^{(1)} - E_i^{(1)}) \psi^{(0)} = 0 \qquad \text{(XXXIII-6)}$$

with

$$E_i^{(1)} = (\psi^{(0)}, H_i^{(1)} \psi^{(0)}) \qquad \text{(XXXIII-7)}$$

Also if we define the $E_{ij}^{(2)\prime}$ by

$$E_{ij}^{(2)\prime} = E_{ij}^{(2)} - (\psi^{(0)}, H_{ij}^{(2)} \psi^{(0)}) \qquad \text{(XXXIII-8)}$$

then one finds that the $E_{ii}^{(2)\prime}$ are given by formulae like (XXXI-11) - (XXXI-13) while the essentially new quantity $E_{12}^{(2)\prime}$ can be written alternatively as

$$E_{12}^{(2)\prime} = \frac{1}{2} \left[(\psi^{(0)}, (H_1^{(1)} - E_1^{(1)}) \psi_2^{(1)}) + c.c. \right] \qquad \text{(XXXIII-9)}$$

or

$$E_{12}^{(2)\prime} = \frac{1}{2} \left[(\psi^{(0)}, (H_2^{(1)} - E_2^{(1)}) \psi_1^{(1)} + c.c. \right] \qquad \text{(XXXIII-10)}$$

or

$$E^{(2)\prime}_{12} = -\tfrac{1}{2}\left[\left(\Psi^{(1)}_1, (H^{(0)} - E^{(0)})\Psi^{(1)}_2\right) + c.c.\right] \qquad \text{(XXXIII-11)}$$

Equations (9) and (10) are, of course, essentially the same as (XXV-54) and (XXV-57). We know how to derive a variational upper bound to the $E^{(2)\prime}_{12}$ when $\Psi^{(0)}$ is the ground state (and in the next section we will discuss a method which also works for excited states), so that our real concern is with approximating $E^{(2)\prime}_{12}$. Further the reader should realize that many physically interesting quantities are of this mixed type: various kinds of shielding factors, spin-spin coupling constants, various optical rotation coefficients, etc.

One method used in practice is to calculate a $\hat{\Psi}^{(1)}_1$, or a $\hat{\Psi}^{(1)}_2$ by use of $\hat{J}^{(2)}_{11H}$ or $\hat{J}^{(2)}_{22H}$ and then to use (9) or (10) with the $\Psi^{(1)}_i$ replaced by $\hat{\Psi}^{(1)}_i$ as approximations. If one calculates both $\hat{\Psi}^{(1)}_1$ and $\hat{\Psi}^{(1)}_2$ then one might also envisage using a formula like (11). However in such a case it is probably better to use the combination (9) + (10) - (11), i.e.,

$$\hat{J}^{(2)}_{12H} \equiv \tfrac{1}{2}\left\{\left(\Psi^{(0)}_1, (H^{(0)} - E^{(0)})\hat{\Psi}^{(1)}_2\right) + \left(\hat{\Psi}^{(1)}_1, (H^{(0)} - E^{(0)})\Psi^{(0)}\right) + \left(\Psi^{(0)}, (H^{(1)}_1 - E^{(1)}_1)\hat{\Psi}^{(1)}_2\right)\right\} + c.c. \qquad \text{(XXXIII-12)}$$

which arises when one makes appropriate replacements in $V^2\hat{J}^{(2)}_H$. $\hat{J}^{(2)}_{12H}$ then has the theoretical virtue that

$$\delta J^{(2)}_{12H} = 0 \text{ all } \delta\Psi^{(1)}_1, \text{ all } \delta\Psi^{(1)}_2$$

so that

$$\hat{J}^{(2)}_{12H} - E^{(2)\prime}_{12}$$

is of order $(\Psi^{(1)}_1 - \hat{\Psi}^{(1)}_1) * (\Psi^{(1)}_2 - \hat{\Psi}^{(1)}_2)$, i.e., it is bilinear in

the errors in $\hat{\Psi}^{(1)}_{1}$ and $\hat{\Psi}^{(1)}_{\sim}$ (see also Eq. (14) below).

However in certain applications, particularly in the calculation of spin-spin coupling constants (for a recent concise review see de Jeu[2]) these procedures are suspect because the conventional $H^{(1)}_{1}$ and $H^{(1)}_{\sim}$ are so singular that $E^{(2)\prime}_{11}$ and $E^{(2)\prime}_{\sim\sim}$ are infinite and so the $J^{(2)}_{11\,H}$ variational problems are really undefined. One can of course introduce cut-offs to make $H^{(1)}_{1}$ and $H^{(1)}_{\sim}$ non singular, and there is some physical basis for doing so. However then one would want to be sure that the approximation to $E^{(2)}_{1\sim}$ calculated in this way was, like $E^{(2)}_{1\sim}$ itself, insensitive to the details of the cut-off procedure.[3]

Another method which has been used is to determine optimal approximations to $\Psi^{(1)}_{1}$ and $\Psi^{(1)}_{2}$ not by use of $\hat{J}^{(2)}_{11\,H}$ or $\hat{J}^{(2)}_{22H}$ but by requiring

$$\delta \check{J}^{(2)}_{12H} = 0 \qquad \text{"all" } \delta\check{\Psi}^{(1)}_{1}\text{, "all" } \delta\check{\Psi}^{(1)}_{2} \qquad \text{(XXXIII-13)}$$

where we have changed notation from $\hat{\ }$ to $\check{\ }$ in order to emphasize that the optimal functions which will emerge from this procedure will in general be different from those derived from $\hat{J}^{(2)}_{11\,H}$ or $\hat{J}^{(2)}_{22H}$. However this procedure can lead to numerical difficulties,[4] though if, in each of the methods, the trial functions are all drawn from the same linear space, it is easy to see that all the methods yield the same results.

A basic difficulty with all these ways of approximating $E^{(2)\prime}_{12}$ is one which we have not yet mentioned; namely even if $\Psi^{(0)}$ is the ground state, none of them yields a bound on $E^{(2)\prime}_{12}$, and therefore

none are capable of systematic improvement. Thus for example, one has

$$\hat{T}^{(2)}_{12H} - E^{(2)\prime}_{12} = \frac{1}{2}\left[(\Delta_1, (H^{(0)} - E^{(0)})\Delta_2) + c.c.\right] \qquad \text{(XXXIII-14)}$$

where

$$\Delta_i = \hat{\psi}^{(1)}_i - \psi^{(1)}_i \qquad \text{(XXXIII-15)}$$

so that in general $(\hat{T}^{(2)}_{12H} - E^{(2)\prime}_{12})$ may have either sign.

One can also see this in another way by noting that if we generalize the argument which led to (XXXI-6) to two perturbations then we find

$$\tilde{E}^{(2)\prime}_{11} + 2\eta\,\tilde{E}^{(2)\prime}_{12} + \eta^2\tilde{E}^{(2)\prime}_{22} \geq E^{(2)\prime}_{11} + 2\eta E^{(2)\prime}_{12} + \eta^2 E^{(2)}_{22} \qquad \text{(XXXIII-16)}$$

where we have written

$$\nu_2 = \eta\nu_1 \qquad \text{(XXXIII-17)}$$

For $\eta = 0$ and $\eta = \infty$ respectively this yields the results we already know, that for the ground state

$$\tilde{E}^{(2)\prime}_{11} \geq E^{(2)\prime}_{11}\,; \quad \tilde{E}^{(2)\prime}_{22} \geq E^{(2)}_{22}$$

but clearly it permits no conclusion about the sign of $\tilde{E}^{(2)\prime}_{12} - E^{(2)\prime}_{12}$. (Thus $(A+B)^2$ is ≥ 0 whatever the relative sign of A and B).

A way out of this difficulty, but one which we will not pursue here,[5] would be to find an improvable <u>lower</u> bound to $E^{(2)}$. Then with the choice

$$H^{(1)} = H^{(1)}_1 + H^{(1)}_2$$

we would be able to calculate upper and lower bounds to

$$E^{\prime}_{+} \equiv E^{(2)\prime}_{11} + 2E^{(2)\prime}_{12} + E^{(2)\prime}_{22} \qquad \text{(XXXIII-18)}$$

while with the choice

$$H^{(1)} = H_1^{(0)} - H_2^{(0)}$$

we could calculate upper and lower bounds to

$$E_-' = E_{11}^{(2)\prime} - 2E_{12}^{(2)\prime} + E_{22}^{(2)\prime} \qquad \text{(XXXIII-19)}$$

Thus we would have

$$E_+^L \leq E_+ \leq E_+^U$$
$$-E_-^U \leq -E_- \leq -E_-^L$$

and therefore by addition of these inequalities

$$E_+^L - E_-^U \leq 4E_{12}^{(2)\prime} \leq E_+^U - E_-^U \qquad \text{(XXXIII-20)}$$

XXXIV. UPPER BOUNDS TO SECOND ORDER ENERGIES FOR EXCITED STATES

In this section we wish to describe two sets of circumstances, really two ways of choosing the $\Psi^{(1)}$, so that $\hat{E}^{(2)\prime}$ will be an upper bound to $E^{(2)}$ when $\psi^{(0)}$ is an excited state. The discussion will closely parallel our earlier discussions of conditions under which an $\hat{E}$ will be an upper bound to an excited state E. Also we will briefly indicate the extension of our results to higher orders.

In Section II we pointed out that we would have $\hat{E} > E$ if $\hat{\psi}$ were orthogonal to all lower states. However we dismissed this as being of no practical importance (in the absence of symmetry considerations), since one would almost certainly not know the lower states exactly. However in the present context the analogous result becomes more interesting. The point is that one may know the exact <u>zero</u> order

wave functions for the lower states and hence can enforce orthogonality through first order. In view of the ν^{2n+1} theorem it is then not surprising that, as we will show in a moment, this will then imply

$$\hat{E}^{(2)'} \rightarrow E^{(2)'} \qquad \text{(XXXIV-1)}$$

Similarly if one knows the lower states exactly through first-order one can ensure that[2]

$$\hat{E}^{(4)'} \rightarrow E^{(4)'} \qquad \text{(XXXIV-2)}$$

etc.

Suppose then that there are T zero order states with normalized eigenfunctions $\psi_t^{(0)}$ and eigenvalues $E_t^{(0)}$ such that $E_t^{(0)} < E^{(0)}$ (more precisely, if there are symmetry considerations involved, T is the number of states with $E_t^{(0)} < E^{(0)}$ and having the symmetry of $\psi^{(0)}$). We now require that

$$\hat{\psi} = \psi^{(0)} + \nu \hat{\psi}^{(1)} + \cdots \qquad \text{(XXXIV-3)}$$

be orthogonal to

$$\psi_t = \psi_t^{(0)} + \nu \psi_t^{(1)} + \cdots \qquad \text{(XXXIV-4)}$$

through first-order. Since they are already orthogonal in zero order, this means that we must have

$$(\psi_t^{(0)}, \hat{\psi}^{(1)}) + (\psi_t^{(1)}, \psi^{(0)}) = 0 \qquad \text{(XXXIV-5)}$$

However from the sum over states formula for $\psi_t^{(1)}$ it follows that

$$(\psi_t^{(1)}, \psi^{(0)}) = - \frac{(\psi_t^{(0)}, H^{(1)}\psi^{(0)})}{E^{(0)} - E_t^{(0)}} \qquad \text{(XXXIV-6)}$$

whence (5) yields

$$(\psi_t^{(0)}, \psi^{(1)}) = - \frac{(\psi_t^{(0)}, H^{(1)}\psi^{(0)})}{E_t^{(0)} - E^{(0)}} \qquad \text{(XXXIV-7)}$$

which, from the sum over states formula for $\psi^{(1)}$ is precisely $(\psi_t^{(0)}, \psi^{(1)})$. Therefore if (5) is satisfied we will have

$$(\psi_t^{(0)}, (\psi^{(1)} - \psi^{(1)})) = 0 \qquad t = 1 \cdots T \qquad \text{(XXXIV-8)}$$

i.e., $\Delta^{(1)}$ of (XXXI-1) is orthogonal to all $\psi_t^{(0)}$ with $E_t^{(0)} < E^{(0)}$ and hence from (XXXIV-8) we will have (1) as announced.

One way[1] to ensure that (5) will be satisfied is simply to use trial functions of the form

$$\tilde{\psi}^{(1)} = - \sum_{t=1}^{T} \psi_t^{(0)} \frac{(\psi_t^{(0)}, H^{(1)}\psi^{(0)})}{E_t^{(0)} - E^{(0)}} + \tilde{X} \qquad \text{(XXXIV-9)}$$

where $\tilde{X}$ is orthogonal to the $\psi_t^{(0)}$. Another more flexible method[3] is to use trial functions of the form

$$\tilde{\psi}^{(1)} = \sum_{t=1}^{T} \tilde{a}_t \psi_t^{(0)} + \tilde{Y} \qquad \text{(XXXIV-10)}$$

where the $\tilde{a}_t$ are linear variation parameters and where $\tilde{Y}$ can be anything except that it should not involve the $\tilde{a}_t$ explicitly.

To see that this works, note that if we ask for the optimal $\tilde{a}_t$ for a given $\tilde{Y}$, let us denote them by $\check{a}_t$, then they will be determined by

$$\delta\tilde{a}_t \left[(\psi_t^{(0)}, (H^{(0)} - E^{(0)})\tilde{\psi}^{(1)} + (H^{(1)} - E^{(1)})\psi^{(0)} \right] = 0 \qquad \text{(XXXIV-11)}$$

or

$$(\psi_t^{(0)}, \tilde{\psi}^{(1)}) = - \frac{(\psi_t^{(0)}, H^{(1)}\psi^{(0)})}{E_t^{(0)} - E^{(0)}} \qquad \text{(XXXIV-12)}$$

and hence (5) will be satisfied even after we determine $\hat{\gamma}$ and thereby turn $\tilde{\psi}^{(1)}$ into $\hat{\psi}^{(1)}$. Note however, that unless $\hat{\gamma}$ is orthogonal to the $\psi_t^{(0)}$

$$\delta_t \neq - (\psi_t^{(0)}, H^{(1)}\psi^{(0)}) / (E_t^{(0)} - E^{(0)}) \qquad \text{(XXXIV-13)}$$

Let us now turn to the perturbation analogue of the results of Sec. XXIX. To make the argument as simple as possible let us in the first instance ignore any possible simplifications due to symmetry so that T denotes the total number of eigenvalues of $H^{(0)}$ below $E^{(0)}$. Then by continuity, this ordering of states will not change as we let ν be different from zero as long as ν is small enough. Suppose then we do a linear variation calculation for the Hamiltonian H, and suppose that the basis set is such that it yields T $\hat{E}_k$ which lies below E (clearly it cannot yield more but it could yield less). Then from the results of Sec. XXIX it follows that

$$\hat{E}_{T+1} \geq E^{(0)} + \nu E^{(1)} + \nu^2 E^{(2)} + \cdots \qquad \text{(XXXIV-14)}$$

Let us now suppose further that $\psi^{(0)}$ is in the linear space. Then clearly we will have

$$\hat{E}_{T+1} = E^{(0)} + \nu E^{(1)} + \nu^2 \hat{E}^{(2)} + \cdots \qquad \text{(XXXIV-15)}$$

and thus, by the standard argument

$$\hat{E}^{(n)} \geq E^{(n)} \tag{XXXIV-16}$$

and therefore

$$\hat{E}^{(n)'} \geq E^{(n)'} \tag{XXXIV-17}$$

as desired. Thus we conclude that if in the Hylleraas variation method we use a linear basis set which includes $\psi^{(0)}$ and has T $\hat{E}_t^{(0)}$ which are less than $E^{(0)}$ then (17) will be satisfied. Similarly if $\psi^{(0)}$ is also in the linear space it will follow that

$$\hat{E}^{(4)} \geq E^{(4)'} \tag{XXXIV-18}$$

etc. Also, of course, the set can contain non-linear parameters so long as the required conditions are met when they take on their optimal values.

In the derivation of (17) we have assumed that $\psi^{(0)}$ was in the linear space. However clearly we may drop this assumption since, as we discussed in Sec. XXXII, given a linear space and then adding $\psi^{(0)}$ to it does not change $\hat{E}^{(n)'}$ in any way. Further if symmetry considerations are involved so that the states $\psi_t^{(0)}$ which can contribute to $E^{(n)'}$ have a definite symmetry then if we restrict the $\tilde{\psi}^{(0)}$ to have that symmetry, T becomes the number of states of that symmetry with $E_t^{(0)} < E^{(0)}$ since adding in functions of other symmetry will also not change $\hat{E}^{(n)'}$. In Appendix C we give another, more detailed, derivation of (17).

If in this second procedure we choose T of the basis functions to be the $\psi_t^{(0)}$ then, of course, we are essentially back to (10) again.

Also, in view of the results found in Sec. B, we will get the greatest numerical accuracy in this way.

XXXV. SPECIAL THEOREMS SATISFIED BY OPTIMAL FIRST-ORDER WAVEFUNCTIONS

We will derive these theorems by a perturbation analysis of various results given in Sec. XIV et seq. We first note that many of the theorems in question were consequences of equations of the form

$$(\Delta\Phi, (H-\hat{E})\Phi) + (\Psi, (H-E)\Delta\Phi) = 0 \qquad \text{(XXXV-1)}$$

or

$$(\Delta\Phi, (H-\hat{E})\Phi) = 0 \qquad \text{(XXXV-2)}$$

for various choices of $\Delta\Phi$.

Making a perturbation analysis of (1) we then find, through second order

$$(\Delta^0\Phi, (H^{(0)}-\hat{E}^{(0)})\Phi^{(0)}) + c.c. = 0 \qquad \text{(XXXV-3)}$$

$$(\Delta^0\Phi, (H^{(0)}-\hat{E}^{(0)})\Phi^{(1)} + (H^{(1)}-\hat{E}^{(1)})\Phi^{(0)}) + (\Delta^1\Phi, (H^{(0)}-\hat{E}^{(0)})\Phi^{(0)}) + c.c. = 0 \qquad \text{(XXXV-4)}$$

and

$$(\Delta^0\Phi, (H^{(0)}-\hat{E}^{(0)})\Phi^{(2)} + (H^{(1)}-\hat{E}^{(1)})\Phi^{(1)} + (H^{(2)}-\hat{E}^{(2)})\Phi^{(0)}) + (\Delta^1\Phi, (H^{(0)}-\hat{E}^{(0)})\Phi^{(1)} + (H^{(1)}-\hat{E}^{(1)})\Phi^{(0)}) + (\Delta^2\Phi, (H^{(0)}-\hat{E}^{(0)})\Phi^{(0)}) + c.c. = 0 \qquad \text{(XXXV-5)}$$

Let us first consider the situation which we have been discussing most recently, namely

$$\Phi^{(0)} = \Psi^{(0)} \quad , \hat{E}^{(0)} = E^{(0)} \quad , \hat{E}^{(1)} = E^{(1)} \qquad \text{(XXXV-6)}$$

Then (3) is satisfied identically, and also the $\Delta' \hat{\Psi}$ part of (4) vanishes identically. Thus the theorem in question will be satisfied through first-order if

$$\left(\Delta^0 \hat{\Psi}, \, (H^{(0)} - E^{(0)}) \Psi^{(1)} + (H^{(1)} - E^{(1)}) \Psi^{(0)} \right) + c.c. = 0 \qquad \text{(XXXV-7)}$$

Comparison with (XXX-14) then tells us that (6) will be satisfied if

$$\delta\epsilon \, \Delta^0 \Psi \qquad \text{(XXXV-8)}$$

with $\delta\epsilon$ real, is a possible variation of $\Psi^{(1)}$ within the set of trial functions. One way to ensure this is, and indeed for several $\Delta^0 \Psi$ simultaneously is clearly to use trial functions of the form

$$\tilde{\Psi}^{(1)} = \sum_i \tilde{\alpha}_i \Delta_i^0 \hat{\Psi} + \tilde{\gamma} \qquad \text{(XXXV-9)}$$

with the $\tilde{\alpha}_i$ variational parameters and $\tilde{\gamma}$ independent of $\tilde{\alpha}$. Indeed with this choice (8) will be a possible variation, also with $\delta\epsilon$ pure imaginary whence (2) will also be satisfied through first-order.

We now consider some special cases:

(i) From the results of Sec. XVI it then follows that the generalized Hellmann-Feynman theorem will be satisfied through first-order in ν if $\delta\epsilon \, \partial\Psi^{(0)}/\partial\sigma$ with $\delta\epsilon$ real is a possible variation of $\hat{\Psi}^{(1)}$.

(ii) The hypervirial theorem for any G will be satisfied through first-order if $i\delta\epsilon G^{(0)} \Psi^{(0)}$ with $\delta\epsilon$ real, is a possible variation of $\Psi^{(1)}$. If $\delta\epsilon$ may also be pure imaginary then (XVII-6) will also be satisfied through first-order.

(iii) The various theorems of Sec. XXII will be satisfied through first-order if $\zeta\,\delta\epsilon\, C^{(0)}\psi_a^{(0)}$ and $\tau\,\delta\epsilon\, C^{(0)}\psi_b^{(0)}$ with $\delta\epsilon$ an infinitesimal but otherwise arbitrary complex number are possible variations of $\hat{\psi}_b^{(1)}$ and $\psi_a^{(1)}$ respectively. In particular, use of trial functions of the form $\hat{\psi}^{(1)} = \hat{\chi} + \xi\,\psi_b^{(0)}$ and $\hat{\psi}_b^{(1)} = \hat{\chi} + \tau_b\,\psi^{(0)}$ will ensure that $\psi^{(0)} + \nu\,\hat{\psi}^{(1)}$ and $\psi_b^{(0)} + \nu\,\hat{\psi}_b^{(1)}$ are orthogonal through first-order in ν. Note that this is different from our earlier result (XXXIV-10) which was that use of such a $\hat{\psi}^{(1)}$ would ensure that $\psi^{(0)} + \nu\,\hat{\psi}^{(1)}$ would be orthogonal through first-order to $\psi_b^{(0)} + \nu\,\psi_b^{(1)}$.

Let us now consider the second order equation (5). Since it involves $\psi^{(2)}$ we can't in general hope to satisfy it knowing only $\psi^{(0)}$ and $\psi^{(1)}$. However if

$$\Delta^0\psi = 0 \qquad \text{(XXXV-9)}$$

(which, incidentally, will satisfy (4) identically) then, since the $\Delta^2\psi$ term in (5) vanishes identically, we see that the theorem will be satisfied through second-order provided that

$$\delta\epsilon\,\Delta^1\psi \qquad \text{(XXXV-10)}$$

with $\delta\epsilon$ real, is a possible variation of $\psi^{(1)}$ within the set of trial functions.

An example of a theorem for which (9) holds is the generalized Hellmann-Feynman theorem when $H^{(0)}$ doesn't involve σ so that $\Delta^0\psi = \partial\psi^{(0)}/\partial\sigma$ is identically zero. Then it follows that

to satisfy the generalized Hellmann-Feynman theorem through second-order it is sufficient to use a set of trial functions which is independent of σ since then

$$\delta\sigma\, \Delta' \tilde{\Phi} = \delta\sigma \frac{\partial \Phi^{(0)}}{\partial \sigma}$$

will be a possible variation of $\Phi^{(0)}$ within the set (recall the analogous discussions given in Sec. XVI).

The unitary invariance theorems of Sec. XIV do not fall into the patterns (1) and (2). The basic fact needed in their derivation was that a set of trial functions $\tilde{\Phi}$ used in the variational method for H will yield the same optimal energies as the set $U^{\dagger}\tilde{\Phi}$ used in the variation method for

$$H^{U} = U^{\dagger} H U \qquad \text{(XXXV-11)}$$

Let us first consider situations in which U is independent of z. Then, as we have seen, if we use trial functions of the form

$$\tilde{\Phi} = \Phi^{(0)} + \nu \tilde{\Phi}^{(1)} + \cdots \qquad \text{(XXXV-12)}$$

in the variation method for H, we find

$$\hat{E} = E^{(0)} + \nu E^{(1)} + \nu^2 \hat{E}^{(2)} + \cdots \qquad \text{(XXXV-13)}$$

Similarly if we use trial functions of the form

$$\tilde{\Phi}^{U} = U^{\dagger} \Phi^{(0)} + \nu \tilde{\Phi}^{U(1)} \qquad \text{(XXXV-14)}$$

in the variation method for H^{U} then we find

$$\hat{E}^{U} = E^{(0)} + \nu E^{(1)} + \nu^2 \hat{E}^{U(2)} + \cdots \qquad \text{(XXXV-15)}$$

since $U^{+}\psi^{(0)}$ is the eigenfunction of

$$H^{U(0)} = U^{+} H^{(0)} U$$

belonging to eigenvalue $E^{(0)}$, and since $E^{U(1)}$ will equal

$$E^{U(1)} = (U^{+}\psi^{(0)}, U^{+}H^{(1)}U\, U^{+}\psi^{(0)}) = E^{(1)}$$

If we now write (13) in the form

$$\tilde{\psi}^{U} = U^{+}(\psi^{(0)} + \nu\, U\tilde{\psi}^{U(1)} + \cdots) \qquad \text{(XXXV-16)}$$

then comparison with (12) makes it clear that if we use the same set of trial functions in $\tilde{T}^{(\nu)}_{H}$ and $\tilde{T}^{U(\nu)}_{H}$, that is if the set $\tilde{\psi}^{(1)}$ is identical to the set $\tilde{\psi}^{U(1)}$, then we will have

$$\tilde{E}^{U(\nu)} = E^{(\nu)}$$

i.e., unitary invariance through second-order, provided that the set is invariant to U so that (16) can be written

$$\psi^{U} = U^{+}(\psi^{(0)} + \nu\,\tilde{\psi}^{(1)} + \cdots) \qquad \text{(XXXV-17)}$$

Next let us consider a case in which U depends on ν, and in particular a gauge transformation with $H^{(0)}$ the field free Hamiltonian. Then U takes the form

$$U = e^{i\nu\Gamma} \qquad \text{(XXXV-18)}$$

so that instead of (13) we will have

$$\tilde{\psi}^{U} = \psi^{(0)} + \nu\,\tilde{\psi}^{U(1)} + \cdots \qquad \text{(XXXV-19)}$$

which, since

$$U = 1 + i\nu\Gamma$$

can be written, correctly through first-order, as

$$\tilde{\Psi}^U = U^+ \left(\psi^{(0)} + \nu \left[\tilde{\Psi}^{\nu(1)} + i\Gamma\psi^{(0)} \right] + \cdots \right) \qquad \text{(XXXV-20)}$$

Comparison with (11) then makes it clear that even if we use the same set of trial functions in $\tilde{\mathcal{F}}^{(2)}_{H}$ and $\tilde{\mathcal{F}}^{(2)U}_{H}$, still we will have gauge invariance of the energy through second-order provided that the set is invariant to

$$\tilde{\Psi}^{(1)} \rightarrow \tilde{\Psi}^{(1)} + i\Gamma\psi^{(0)} \qquad \text{(XXXV-21)}$$

For any particular Γ this can of course be assured by using trial functions of the form

$$\tilde{\Psi}^{(1)} = i\tilde{\alpha}\Gamma\psi^{(0)} + \tilde{\chi} \qquad \text{(XXXV-22)}$$

Let us now turn our attention to the variation method within the variation method discussed in Sec. XXVI. In this case (6) no longer holds in general, however we will assume that $\Psi^{(0)}$ does satisfy the zero order theorem (3). Then comparing (4) with (XXVI-11) and (XXVI-8) one sees that (4) will be satisfied with $\Psi^{(1)}$ replaced by $\chi^{(1)}$ provided that

$$\delta\epsilon\, \Delta^0 \Psi \qquad \text{(XXXV-23)}$$

with $\delta\epsilon$ real is a possible $\delta\tilde{\chi}^{(1)}$, and provided that

$$\delta\epsilon\, \Delta^1 \Psi \qquad \text{(XXXV-24)}$$

is a possible $\delta\tilde{\chi}^{(2)'}$, or more generally, is a possible $\delta\tilde{\chi}^{(2)'}$ plus possible $\delta^0\Psi$ since the contribution of the latter will vanish

identically. As an example we will now show that in a variational approximation to UHF, the hypervirial theorem for a one-electron G

$$G = \sum_{s=1}^{N} g(s) \tag{XXXV-25}$$

will be satisfied through first-order if for each K

$$i\,\delta G\, g^{(0)}\,\hat{\varphi}_K^{(0)} \tag{XXXV-26}$$

with δG real is a possible variation of $\overset{\vee}{\varphi}{}_K^{(1)}$.

Since in UHF there is no explicit ν dependence we have (recall also (XXIX-9) and (XXIX-13)) that

$$\hat{\Psi}^{(0)} = |\hat{\varphi}_1^{(0)} \cdots \hat{\varphi}_N^{(0)}| \tag{XXXV-27}$$

$$\Psi^{(1)} = \sum_i |\hat{\varphi}_1^{(0)} \cdots \overset{\vee}{\varphi}{}_i^{(1)} \cdots \hat{\varphi}_N^{(0)}| \tag{XXXV-28}$$

$$\Psi^{(2)'} = \sum_{i\neq j}\sum |\hat{\varphi}_1^{(0)} \cdots \overset{\vee}{\varphi}{}_i^{(1)} \cdots \overset{\vee}{\varphi}{}_j^{(1)} \cdots \hat{\varphi}_N^{(0)}| \tag{XXXV-29}$$

Thus with

$$\Delta\overset{\vee}{\Psi} = i\,G\,\overset{\vee}{\Psi} \tag{XXXV-30}$$

we have

$$\Delta^0\overset{\vee}{\Psi} = i\,G^{(0)}\hat{\Psi}^{(0)} = i\sum_j |\hat{\varphi}_1^{(0)} \cdots g^{(0)}\hat{\varphi}_j^{(0)} \cdots \hat{\varphi}_N^{(0)}| \tag{XXXV-31}$$

and

$$\Delta^1\overset{\vee}{\Psi} = i\,G^{(0)}\overset{\vee}{\Psi}{}^{(1)} + i\,G^{(1)}\hat{\Psi}^{(0)} = i\sum_{i\neq j}\sum \{|\hat{\varphi}_1^{(0)} \cdots g^{(0)}\hat{\varphi}_i^{(0)} \cdots \overset{\vee}{\varphi}{}_j^{(1)} \cdots \varphi_N^{(0)}| + (i\leftrightarrow j)\} + i\sum_j |\hat{\varphi}_1^{(0)} \cdots g^{(0)}\overset{\vee}{\varphi}{}_j^{(1)} \cdots \hat{\varphi}_N^{(0)}| + i\,G^{(1)}\hat{\Psi}^{(0)} \tag{XXXV-32}$$

On the other hand from (29) and (30) we also have

$$\delta\Psi^{(1)} = \sum_i |\hat{\varphi}_1^{(0)} \cdots \delta\overset{\vee}{\varphi}{}_i^{(1)} \cdots \hat{\varphi}_N^{(0)}| \tag{XXXV-33}$$

and

$$\delta\check{\Psi}^{(n)\prime} = \sum_{i\neq j}\sum \left\{ \left| \hat{\varphi}_1^{(0)} \cdots \delta\check{\varphi}_i^{(1)} \cdots \check{\varphi}_j^{(1)} \cdots \hat{\varphi}_N^{(0)} \right| + (i \leftrightarrow j) \right\} \qquad \text{(XXXV-34)}$$

Thus if (26) is a possible $\delta\check{\varphi}_k^{(1)}$ we see that, as desired, $\delta\epsilon\, \Delta^0\check{\Psi}$ will be a possible $\delta\check{\Psi}^{(0)}$ and the double sum in $\delta\epsilon\, \Delta^1\check{\Psi}$ will be a possible $\delta\check{\Psi}^{(n)\prime}$. Then, since the single sum in $\delta\epsilon\, \Delta^1\hat{\Psi}$, and also $\sum_{(k)} \hat{\Psi}^{(0)}$ are clearly possible $\delta^0\Phi$ (with $\delta\hat{\varphi}_i^{(0)} = \epsilon\delta\epsilon\, g^{(0)}\hat{\varphi}_i^{(0)}$ and $\delta\hat{\varphi}_i^{(0)} = \epsilon\delta\epsilon\, g^{(0)}\hat{\varphi}_i^{(0)}$ respectively) we have the announced result.

XXXVI. THE DELVES VARIATION METHODS

In Sec. XXIV and in Sec. XXXI we mentioned the use of approximations to $\Psi^{(0)}$ and of model Hamiltonians; the inaccuracies then to be corrected by use of perturbation theory. We will now discuss such procedures in more detail. Typically one has a normalized function $\chi^{(0)}$ which in some sense is an approximation to an eigenfunction $\Psi^{(0)}$ of an isolated system. Then to approximate the effects of external perturbations[A]

$$v_1 H_1^{(1)} + v_1^2 H_1^{(2)} + \cdots \qquad \text{(XXXVI-1)}$$

on the actual system one finds a "model Hamiltonian" $H^{(00)}$ of which $\chi^{(0)}$ is an eigenfunction

$$(H^{(00)} - E^{(00)})\chi^{(0)} = 0 \qquad \text{(XXXVI-2)}$$

and then tries to find the eigenvalues and eigenfunctions of

$$H' = H^{(00)} + v_1 H_1^{(1)} + v_1^2 H_1^{(2)} + \cdots \qquad \text{(XXXVI-3)}$$

using, for example, the perturbation and perturbation variation procedures which we have just been discussing. The results are then taken to approximate the behavior of the actual system. Further, having done this, one can now go on and correct these results by introducing $H^{(0)} - H^{(00)}$ as another perturbation

$$(H^{(0)} - H^{(00)}) \equiv \nu_2 H^{(1)}_2 \qquad \text{(XXXVI-3)}$$

possibly using interchange theorems to simplify the calculations if, as we will assume, there are no problems with degeneracy.

In particular, since we are dealing with exact eigenfunctions of various Hamiltonians, appropriate Hellmann-Feynman theorems will all be valid so that we can use formulae like (XXV-57) or (XXXIII-10) and (XXV-60) to determine the first-order corrections. Thus denoting the appropriate eigenfunction of (2) by Φ and the eigenvalue by Ξ and writing

$$\Phi = \chi^{(0)} + \nu_1 \Phi^{(1)} + \nu_1^2 \Phi^{(2)} + \cdots \qquad \text{(XXXVI-4)}$$

then we have

$$\Xi^{(2)}_{12} = \frac{1}{2}\left\{ -\left(\chi^{(0)}, (H^{(1)}_2 - \langle H^{(1)}_2 \rangle_0)\Phi^{(1)}\right) + c.c. \right\} \qquad \text{(XXXVI-5)}$$

and

$$\Xi^{(3)}_{112} = \frac{1}{3}\left\{ \left(\Phi^{(1)}, (H^{(1)}_2 - \langle H^{(1)}_2 \rangle_0)\Phi^{(1)}\right) + \left(\Phi^{(2)}, (H^{(1)}_2 - \langle H^{(1)}_2 \rangle_0)\chi^{(0)}\right) + c.c. \right\} \qquad \text{(XXXVI-6)}$$

where

$$\langle H^{(1)}_2 \rangle_0 = (\chi^{(0)}, H^{(1)}_2 \chi^{(0)}) \qquad \text{(XXXVI-7)}$$

Further, from (5), the average value of $H_1^{(1)}$ for $\nu_1=0$, correct through first-order in ν_2 is

$$\left(\frac{\partial E}{\partial \nu_1}\right)_{\nu_1=0} = \left(\frac{\partial E}{\partial \nu_1}\right)_{\substack{\nu_1=0\\ \nu_2=0}} + 2\nu_2 E^{(2)}_{12} + \cdots$$

which, using (2) and (3) can be written

$$\left(\frac{\partial E}{\partial \nu_1}\right)_{\nu_1=0} = \langle H_1^{(1)}\rangle_0 + \left\{\left(\chi^{(0)}, (H^{(0)} - \langle H^{(0)}\rangle_0)\phi^{(1)}\right) + \text{c.c.}\right\} \quad \text{(XXXVI-8)}$$

Similarly from (6) the "polarizability" $\left(\partial^2 E/\partial \nu_1^2\right)_{\nu_1=0}$ is given correctly through first-order in ν_2 by

$$\left(\frac{\partial^2 E}{\partial \nu_1^2}\right)_{\nu_1=0} = \left(\frac{\partial^2 E}{\partial \nu_1^2}\right)_{\substack{\nu_1=0\\ \nu_2=0}} + 3\nu_2 E^{(3)}_{112} + \cdots$$

$$= \left(\frac{\partial^2 E}{\partial \nu_1^2}\right)_{\substack{\nu_1=0\\ \nu_2=0}} + \nu_2\Big[\left(\phi^{(1)}, (H_2^{(1)} - \langle H_2^{(1)}\rangle_0)\phi^{(1)}\right)$$

$$+ \left(\phi^{(2)}, (H_2^{(0)} - \langle H_2^{(0)}\rangle_0)\chi^{(0)}\right) + \left(\chi^{(0)}, (H_2^{(0)} - \langle H_2^{(0)}\rangle_0)\phi^{(2)}\right)\Big] + \cdots \quad \text{(XXXVI-9)}$$

which, using (2), (3), and the equation for $\phi^{(1)}$

$$(H^{(00)} - E^{(00)})\phi^{(1)} + (H_1^{(1)} - \langle H_1^{(1)}\rangle_0)\chi^{(0)} = 0 \quad \text{(XXXVI-10)}$$

can be written[B]

$$\left(\frac{\partial^2 E}{\partial \nu_1^2}\right)_{\nu_1=0} = \left(\frac{\partial^2 E}{\partial \nu_1^2}\right)_{\substack{\nu_1=0\\ \nu_2=0}} + \Big\{\left(\phi^{(1)}, (H^{(0)} - \langle H^{(0)}\rangle_0)\phi^{(1)}\right) +$$

$$+ \left(\phi^{(1)}, (H_1^{(1)} - \langle H_1^{(1)}\rangle_0)\chi^{(0)}\right) \quad \text{(XXXVI-11)}$$

$$+ \left(\phi^{(2)}, (H^{(0)} - \langle H^{(0)}\rangle_0)\chi^{(0)}\right) + \left(\chi^{(0)}, (H^{(0)} - \langle H^{(0)}\rangle_0)\phi^{(2)}\right)\Big\} + \cdots$$

The procedure which we have just sketched is actually rather a formal one since (recall the analogous discussion in Sec. XI) there are many choices for $H^{(00)}$ which will satisfy (2), and in general different choices for $H^{(00)}$ will, for example, yield different answers for "the" first-order corrections (5) and (6). On occasion there may be a natural choice for $H^{(00)}$; however usually this will not be the case.

We now note, following Delves,[1] that the variation method suggests another approach to calculating first-order corrections which seems to avoid the problem of choosing an $H^{(00)}$. We will confine our attention to first-order corrections to expectation values; the generalization to polarizations will be obvious.

The actual Hamiltonian of the system is

$$H = H^{(0)} + \nu_1 H_1^{(1)} + \nu_1^2 H_1^{(2)} + \cdots$$

Consider now the quantity $\mathcal{E}$ defined by

$$\mathcal{E} \equiv (\chi, H\chi)/(\chi,\chi) \qquad \text{(XXXVI-12)}$$

where χ , much like $\tilde{\psi}$ in Eq. (XXVI-41) et seq., is some function which, for $\nu_1 = 0$, becomes equal to $\chi^{(0)}$.

Then since $(\psi - \chi)$ will in general involve terms of order ν_1 and of order γ where γ is some measure of $\psi^{(0)} - \chi^{(0)}$ it follows that $(\mathcal{E} - \mathcal{E})$ will involve terms of order ν_1^2 , γ^2 and $\nu_1\gamma$, and hence $\frac{\partial}{\partial \nu_1}(\mathcal{E} - \mathcal{E})$ will be linear in γ .

However now suppose that χ is such that if γ were zero then χ would be exact. Under this circumstance $(\psi - \chi)$

will not involve a term linear in γ and hence we will have that

$$\frac{\partial}{\partial \nu_1}\,(E-\epsilon)\Big|_{\nu_1=0} = 0(\gamma^2)$$

i.e.

$$\left(\frac{\partial \epsilon}{\partial \nu_1}\right)_{\nu_1=0} \equiv \epsilon^{(1)}$$

will yield an approximation to the expectation value of $H_1^{(1)}$ for $\nu_1=0$ which has only a second order error.

This result is similar to that which we found in Sec. XVI. They become identical if, as we assumed there, χ also satisfies the generalized Hellmann-Feynman theorem for $\sigma = \nu_1$ and the 'Hamiltonian' H since then

$$\left(\frac{\partial \epsilon}{\partial \nu_1}\right)_{\nu_1=0} = \langle H_1^{(0)} \rangle_0$$

and we have the result found in Sec. XVI that under the given conditions $\langle H_1^{(1)} \rangle_0$ is already accurate through first order.

Returning to the general case and expanding in powers of ν, we find from (10) or from (XXVI-43) and (XXVI-44) that

$$\epsilon^{(0)} = (\chi^{(0)}, H^{(0)}\chi^{(0)}) = \langle H^{(0)} \rangle_0 \qquad \text{(XXXVI-13)}$$

and

$$\epsilon^{(1)} = \langle H_1^{(1)} \rangle_0 + \{(\chi^{(0)}, (H^{(0)} - \epsilon^{(0)})\chi^{(1)}) + c.c.\} \qquad \text{(XXXVI-14)}$$

Comparison of (14) with (8) then shows that we will formally recover our perturbation results if we choose $\chi^{(1)}$ equal to $\phi^{(1)}$. However, as we have said, $\phi^{(1)}$ is not really defined until one chooses an $H^{(00)}$. The suggestion of Delves, which seems to bypass this ambiguity, is to note that there is a very natural choice

for $\chi^{(0)}$, namely it should be the solution of the equation

$$(H^{(0)} - \epsilon^{(0)})\chi_D^{(1)} + (H_1^{(1)} - \mathcal{J})\chi^{(0)} = 0 \qquad \text{(XXXVI-15)}$$

where $\mathcal{J}$ is to be chosen so that

$$(\chi^{(0)}, \chi_D^{(1)}) = 0 \qquad \text{(XXXVI-16)}$$

This, of course, is just the equation we wrote down earlier (in Sec. XXXI) in a somewhat different notation and the point is that, as we discussed in Sec. XXXI, if $\chi^{(0)} \to \psi^{(0)}$ then $\chi_D^{(1)}$ will tend smoothly to the exact result $\psi_1^{(1)}$,as desired.

Also, from (15) we have

$$\mathcal{J} = \langle H_1^{(1)} \rangle_0 + (\chi^{(0)}, (H^{(0)} - \epsilon^{(0)})\chi_D^{(1)}) \qquad \text{(XXXVI-17}$$

so that Eq. (14) yields

$$\epsilon^{(1)} = \mathcal{J} - \langle H_1^{(1)} \rangle_0 \qquad \text{(XXXVI-18)}$$

Now an obvious objection to the use of (15) is that one presumably can not solve it exactly and will have to resort to variational approximations. However similar considerations usually apply to the perturbation approach. That is for some special choices of $H^{(00)}$ one may be able to solve

$$(H^{(00)} - \epsilon^{(00)})\phi^{(1)} + (H_1^{(1)} - \langle H_1^{(1)} \rangle_0)\chi^{(0)} = 0 \qquad \text{(XXXVI-19)}$$

however usually we will have to use variation methods to approximate the solution. For details of such calculations we refer the reader to the

literature[1,2,3]. The appropriate variation method for (19) is of course,

$$\delta \hat{\mathcal{J}} = 0 \qquad \text{"all" } \delta \hat{\phi}^{(1)} \tag{XXXVI-20}$$

where

$$\hat{\mathcal{J}} = (\hat{\phi}^{(1)}, (H^{(00)} - \epsilon^{(00)})\hat{\phi}^{(1)}) + (\hat{\phi}^{(1)}, (H^{(1)}_1 - \langle H^{(1)}_1 \rangle_0)\chi^{(0)}) + (\chi^{(0)}, (H^{(1)}_1 - \langle H^{(1)}_1 \rangle_0)\hat{\phi}^{(1)}) \tag{XXXVI-21}$$

while that for (15) is the "Delves variation method"[4]

$$\delta \hat{\mathcal{J}}_D = 0 \qquad \text{"all" } \delta \hat{\chi}^{(1)}_D \tag{XXXVI-22}$$

where

$$\hat{\mathcal{J}}_D = (\hat{\chi}^{(1)}_D, (H^{(00)} - \epsilon^{(00)})\hat{\chi}^{(1)}_D) + (\hat{\chi}^{(1)}_D, (H^{(1)}_1 - \xi)\chi^{(0)}) + (\chi^{(0)}, (H^{(1)}_1 - \xi)\hat{\chi}^{(1)}_D) \tag{XXXVI-23}$$

and where ξ is to be determined so that

$$(\chi^{(0)}, \hat{\chi}^{(1)}_D) = 0$$

Having come this far we must now point out that this (Delves) method can be viewed as a special case of our earlier $H^{(00)}$ approach. That is as is often the case[D] the fact that the symbol "$H^{(00)}$" does not appear explicitly does not necessarily mean that the problem of choosing an $H^{(00)}$ has been avoided; it may mean only that a special choice has been made.

To see this let P be the projector onto $\chi^{(0)}$, i.e., in Dirac notation

$$P = |\chi^{(0)}\rangle\langle\chi^{(0)}|$$

and let us take for $H^{(00)}$

$$H^{(00)} = P H^{(0)} P + (1-P) H^{(0)} (1-P) \qquad \text{(XXXVI-24)}$$

or

$$H^{(00)} = \epsilon^{(0)} P + (1-P) H^{(0)} (1-P) \qquad \text{(XXXVI-25)}$$

Then since from (25)

$$(H^{(00)} - \epsilon^{(0)}) \chi^{(0)} = 0 \qquad \text{(XXXVI-26)}$$

we see that

$$E^{(00)} = \epsilon^{(0)}$$

so that Eq. (19) for $\phi^{(1)}$ becomes

$$(\epsilon^{(0)} P + (1-P) H^{(0)} (1-P) - \epsilon^{(0)}) \phi^{(1)} + (H_1^{(1)} - \langle H_1^{(1)} \rangle_0) \chi^{(0)} = 0 \qquad \text{(XXXVI-27)}$$

Now equation (25) does not determine ($\chi^{(0)}$, $\phi^{(1)}$) and indeed its value is physically irrelevant. Suppose now that we choose it equal to zero. Then we will have

$$P \phi^{(1)} = 0 \qquad \text{(XXXVI-28)}$$

whence since in any case

$$P (H_1^{(1)} - \langle H_1^{(1)} \rangle_0) \chi^{(0)} = 0$$

we can write (27) as

$$(1-P) [(H^{(0)} - \epsilon^{(0)}) \phi^{(1)} + (H^{(1)} - \langle H_1^{(1)} \rangle_0) \chi^{(0)}] = 0$$

or

$$(H^{(0)} - \epsilon^{(0)}) \phi^{(1)} + (H^{(1)} - \langle H_1^{(1)} \rangle_0) \chi^{(0)} + \alpha \chi^{(0)} \qquad \text{(XXXVI-29)}$$

where α is to be chosen so that (28) is satisfied.

However with this condition; we see, comparing (29) with (15) that $\phi^{(1)} = \chi_D^{(1)}$ and thus we have the result that this Delves variation method is a special case of the perturbation method with $H^{(00)}$ given by (24). A natural question at this point is then, is this choice of $H^{(00)}$ in some sense a best choice? We will not attempt to answer this question though some of the material which follows would be relevant to such an attempt.

As implied in our use of the plural (m) the title of this section, the method which we have just been discussing is not the only one which Delves has proposed. We now want to consider a more general formalism which he introduced earlier,[5] and which has recently been employed in various ways by other authors.[2,6]

We have seen that $\epsilon^{(1)}$ defined by (14), which we will now write in the form,

$$0 = \left(\chi^{(0)}, (H^{(1)} - \epsilon^{(1)}) \chi^{(0)} \right) + \left(\chi^{(0)}, (H^{(0)} - \epsilon^{(0)}) \chi^{(1)} \right) + \left(\chi^{(1)}, (H^{(0)} - \epsilon^{(0)}) \chi^{(0)} \right) \quad \text{(XXXVI-30)}$$

in order to free it from the restriction that $\chi^{(0)}$ be normalized, differs from $\left(\frac{\partial E}{\partial \nu_1}\right)_{\nu_1 = 0}$ by terms bilinear in $(\psi^{(0)} - \chi^{(0)})$ and $(\psi_1^{(1)} - \chi^{(1)})$. That is, and this is the Delves variation principle,

$$\delta \epsilon^{(1)} = 0 \quad \text{all } \delta\chi^{(0)} \text{ all } \delta\chi^{(1)} \quad \text{(XXXVI-31)}$$

This observation then suggests using

$$\delta \hat{\epsilon}^{(1)} = 0 \quad \text{"all" } \delta\hat{\chi}^{(1)} \quad \text{(XXXVI-32)}$$

perhaps supplemented by the requirement

$$(\chi^{(0)}, \hat{\chi}^{(k)}) = 0 \qquad \text{(XXXVI-33)}$$

as a means for determining an optimal $\hat{\chi}^{(k)}$ for a given $\chi^{(0)}$.

More generally one might consider requiring

$$\delta \hat{E}^{(1)} = 0 \quad \text{"all" } \delta\hat{\chi}^{(k)} \text{ and "all" } \delta\hat{\chi}^{(0)} \qquad \text{(XXXVI-34)}$$

in order to determine an optimal $\hat{\chi}^{(k)}$ and an optimal $\hat{\chi}^{(0)}$, both adapted to the particular perturbation $H_1^{(1)}$.

However in practice these procedures turn out not to be very satisfactory; for example (32) evidently leads us to an equation like (XXVI-48) which we discussed earlier. The basic difficulty is that for the $\hat{\chi}^{(k)}$ which emerge, $(\psi_1^{(1)} - \hat{\chi}^{(k)})$ may be of zero order in $(\psi^{(0)} - \chi^{(0)})$ and hence $\left(\frac{\partial E}{\partial v_1}\right)_{v_1=0} = \hat{E}^{(1)}$ will be of first-order in $(\psi^{(0)} - \chi^{(0)})$ rather than of second-order as desired. It was for this reason then that Delves suggested that $\chi^{(k)}$ should be determined instead from (15), thereby ensuring that $(\psi_1^{(1)} - \chi^{(1)})$ will be of order $(\psi^{(0)} - \chi^{(0)})$.

Recently however other authors have found other uses for (33) and (34). For example, Shustek and Kreiger[2] have pointed out that one can use (33) to derive the Shielding Approximation mentioned in Sec. XXIX in connection with UHF. Here, proceeding in a somewhat different order than before, one starts with

$$H^{(00)} = T - \sum_{i=1}^{N} \frac{(z_i - s)}{r_i} \qquad \text{(XXXVI-35)}$$

as a model Hamiltonian for an atom of nuclear charge $\mathcal{Z}$ where T is the kinetic energy operator and where the value of the constant S is not specified initially. Equation (8) can then be solved analytically to yield $\chi^{(0)}(S)$ and for many interesting one-electron perturbations Eq. (19) can also be solved analytically to yield $\phi^{(k)}(S)$. Then in (33) we choose

$$\hat{\chi}^{(0)} = \chi^{(0)}(\hat{S}) \qquad \hat{\phi}^{(k)} = \phi^{(k)}(\hat{S}) \tag{XXXVI-36}$$

and determine $\hat{S}$ from

$$\delta \hat{\epsilon}^{(k)} = 0 \qquad \text{all } \delta\hat{S} \tag{XXXVI-37}$$

In many cases it turns out that this is equivalent to requiring that the first-order correction vanish, i.e.,

$$(\chi^{(0)}(\hat{S}),\ (H^{(1)} - \hat{\epsilon}^{(1)})\,\chi^{(0)}(\hat{S})) = 0 \tag{XXXVI-38}$$

which is the prescription of the Shielding Approximation.[E]

In a related vein Kirtman[4] has suggested the following rather general procedure: (i) For a given $\chi^{(0)}$ solve (analytically if possible) Eq. (19), for a set of $(H^{(0)} - \epsilon^{(0)})$ parametrized by some quantities B . All the $(H^{(0)} - \epsilon^{(0)})$ of course should satisfy (2). Then determine $\hat{B}$ from

$$\delta \hat{\epsilon}^{(1)} = 0 \qquad \text{all } \delta\hat{B} \tag{XXXVI-39}$$

where

$$\hat{\chi}^{(1)} = \phi^{(1)}(\hat{B}) \tag{XXXVI-40}$$

(ii) Suppose that $\chi^{(0)}$ involves some quantities C. Then choose an $(\psi^{(00)}, \epsilon^{(00)})$ and solve (19), analytically if possible. Then determine $\hat{C}$ from

$$\delta \hat{\epsilon}^{(1)} = 0 \quad \text{all} \quad \delta \hat{C} \tag{XXXVI-41}$$

where

$$\hat{\chi}^{(0)} = \chi^{(0)}(\hat{C}) \quad , \quad \hat{\chi}^{(1)} = \phi^{(1)}(\hat{C}) \tag{XXXVI-42}$$

Evidently Eq. (36) and (37) are a special case of (41) and (42) however Kirtman usually imposes the additional constraint that $\hat{\epsilon}^{(0)}(\hat{C})$ be fixed at some "reasonable" value so that, for example for him $\hat{C}$ could not be just a single parameter.

APPENDIX C. A DETAILED DERIVATION OF EQ. (XXXIV-17)

Since, as discussed in Sec. XXXII, one can assume without loss of generality that $\psi^{(0)}$ is not in the space, let us do so. Thus we start with (XXXII-12) and, in the standard way, add one more function ϕ to the space. Thus we consider a new set of trial functions

$$\tilde{\chi}^{(0)} = \sum_{k=1}^{M} \tilde{b}_k \dot{\psi}_k^{(0)} + \tilde{b}\,\phi \tag{C-1}$$

where clearly we may assume, without loss of generality, that ϕ is normalized and orthogonal to the $\dot{\psi}_k^{(0)}$

$$(\phi, \phi) = 1 \qquad (\phi, \dot{\psi}_k^{(0)}) = 0 \tag{C-2}$$

One then readily finds that if none of the $\dot{E}_k^{(0)}$ yields $\tilde{E}^{(0)}$ then

$$\delta \tilde{J}_H^{(n)} = 0 \quad \text{all } \delta\tilde{b}_k, \text{ all } \delta\tilde{b} \tag{C-3}$$

yields

$$\tilde{J}_H^{(n)} - \tilde{E}^{(n)\prime} = -\frac{Z}{D(E_0)} \tag{C-4}$$

where $\tilde{E}^{(n)\prime}$ is given by (XXXII-12) and where Z and D are defined as follows:

$$Z = \left| \sum_{k=1}^{M} \frac{(\psi^{(0)}, (H^{(n)} - E^{(n)})\dot{\psi}_k^{(0)})(\dot{\psi}_k^{(0)}, (H^{(0)} - E^{(0)})\phi)}{E^{(0)} - \dot{E}_k^{(0)}} + (\psi^{(0)}, (H^{(n)} - E^{(n)})\phi) \right|^2 \tag{C-5}$$

and is obviously non-negative.

$$D(E^{(0)}) = \sum_k \frac{|(\phi, H^{(0)}\dot{\psi}_k^{(0)})|^2}{E^{(0)} - \dot{E}_k^{(0)}} - \left[E^{(0)} - (\phi, H^{(0)}\phi)\right] \tag{C-6}$$

$$\equiv F(E^{(0)}) - G(E^{(0)}) \tag{C-7}$$

We will now show that if there are T eigenvalues of $H^{(0)}$ below $E^{(0)}$ (whose eigenfunctions have the symmetry of the $\Phi_k^{(0)}$) and if there are T of the $\dot{E}_k^{(0)}$ which are also less than $E^{(0)}$ then $D(E^{(0)})$ is positive. This result then tells us that as we increase the size of the basis set beyond the point where T of the $\dot{E}_k^{(0)}$ are below $E^{(0)}$, $\hat{J}_H^{(0)}$ will then steadily decrease monotonically which, of course, implies (XXIV-17).

The proof follows from the observation that, as one can readily verify, the solutions of

$$D(\mathcal{E}) = 0 \tag{C-8}$$

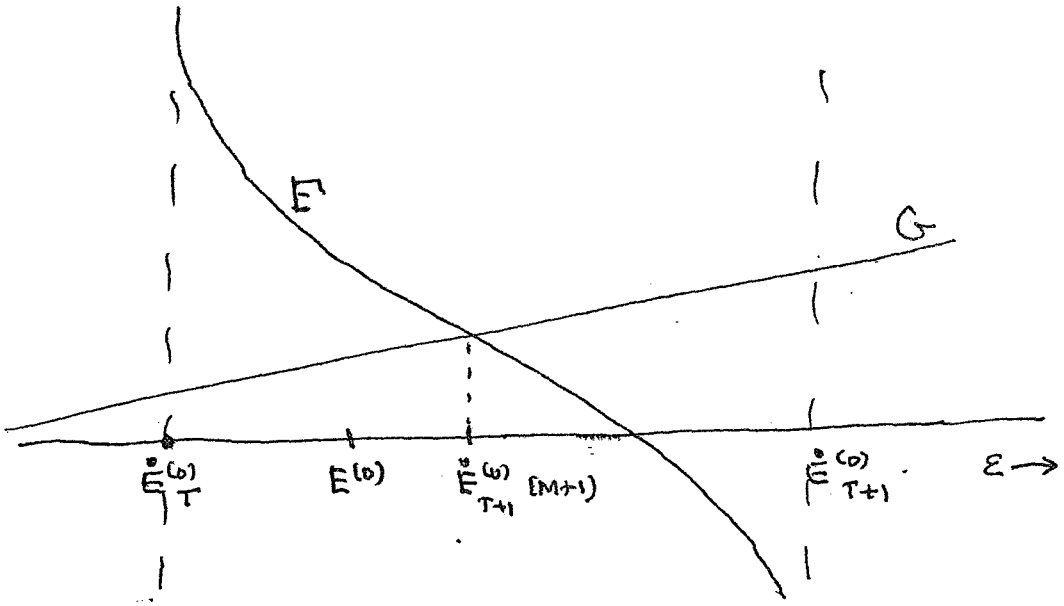
are the $\dot{E}_k^{(0)}(M+1)$ which result by using a trial function of the form (1) in a linear variation calculation for $H^{(0)}$. Thus it follows from (VII-11) that

$$\dot{E}_k^{(0)}(M+1) \leq \dot{E}_k^{(0)} \tag{C-9}$$

Also, from our assumptions

$$\dot{E}_T^{(0)} < E^{(0)} < \dot{E}_{T+1}^{(0)}(M+1) \tag{C-10}$$

Then since $\dot{E}_{T+1}^{(0)}(M+1)$ is given by the appropriate intersection of $F(\mathcal{E})$ and $G(\mathcal{E})$, the positivity of $D(E^{(0)})$ follows just by looking at a graph:



Therefore $F(E^{(\sigma)}) > G(E^{(\sigma)})$ which proves the point.

Clearly, also, we can, by the same sort of argument[2] generalize (XXXII-19) to the case in which Ω has T negative eigenvalues but still no eigenvalue equal to zero, namely again

$$\Omega^{-1} = \sum_{K}\sum_{L} |\phi_L\rangle (e\Omega e^{-1})_{KL} \langle\phi_L| \qquad \text{(C-11)}$$

but now with the additional restriction that the matrix $e\Omega e$ also have T negative eigenvalues. (Notice that this will also ensure that $e\Omega e^{-1}$ will exist).

REFERENCES

SEC. XXIII

1. E. Wigner, Math. Natur. Anz. (Budapest) 53, 477 (1935); R. A. Silverman, Phys. Rev. 85, 227 (1952).
2. See for example, T. Kikuta, Prog. Theo. Phys. 12, 10 (1954); 14, 457 (1955); M. D. Kostin and H. Brooks, J. Math. Phys. 5, 169 (1954). J.O. Hirschfelder J. Chem Phys 39 2009 (1963); R.J. Turnsky, J. Math Anal & Appl. 33 605 (1971).
3. A general discussion has been given by G. C. Pomraning, J. Math. Phys. 8, 149 (1967). This paper also contains references to the earlier literature.
4. W. Kohn, Phys. Rev. 74, 1763 (1948).
5. K. R. Brownstein and W. A. McKinley, Phys. Rev. 170, 1255 (1968), also H. J. Kolker, J. Chem. Phys. 53, 4697 (1970).

SEC. XXIV

1. See for example, H. D. Cohen and C. C. J. Roothaan, J. Chem. Phys. 43, S 34 (1965); A. D. McLean and M. Yoshimine, J. Chem. Phys. 46, 3682 (1967); M. N. Grasso, K. I. Chung, and R. P. Hurst, Phys. Rev. 167, 1 (1968); J. A. Pople, J. W. McIver, Jr., and N. S. Ostlund, J. Chem. Phys. 49, 2960 (1968).
2. See, for example, D. M. Bishop and A. Macias, J. Chem. Phys. 53, 3515 (1970) and references there.
3. For some discussion of a variational approach to degenerate Rayleigh-Schroedinger perturbation theory see S. T. Epstein, J. Chem. Phys. 44, 3266 (1966).

SEC. XXV

1. This case, with the further requirement of no explicit ν dependence, i.e. $\hat{\Phi} = \Theta(\hat{A})$, has been studied extensively by J. N. Silverman and J. L. VanLeuven, Phys. Rev. 162, 1175 (1967). J. E. Harriman, J. Chem. Phys. 54, 902 (1971) discusses the same situation in terms of density matrices.
2. J. N. Silverman, and J. L. Van Leuvan, Chem. Phys. Lett. 7, 37 (1970); (Erratum) 7, 640 (1970).
3. Interchange theorems within UHF have been discussed and applied in many papers by Dalgarno and co-workers; (See Sec. XXIX). See also J. D. Lyons, P. W. Langhoff, and R. P. Hurst, Phys. Rev. 5, 60 (1966).
4. For a discussion of the situation in Rayleigh-Schroedinger perturbation theory (i.e. $\hat{\Phi} = \Psi$) see J. O. Hirschfelder, W. B. Brown, and S. T. Epstein in Advances in Quantum Chemistry 1 (1964), Sec. IVE; P.-O. Löwdin (ed) (Academic Press, New York).

SEC. XXVI

1. For the special cases of UHF and AUHF this has been shown by R. Moccia, Chem. Phys. Lett. 5, 260 (1970); 5, 265 (1970).

SEC. XXVII

1. Within the context of Ref. 1 (OF) Sec. XXV this theorem is proven in the first reference there. For UHF it has in essence been proven by T. K. Rebane, Opt. Spec. 19, 179 (1965); and by P. W. Langhoff, J. D. Lyons, and R. P. Hurst, Phys. Rev. 148, 18 (1966). For AUHF and AOHF many authors have noted that one can calculate $\hat{E}^{(2)}$ and $\hat{E}^{(3)}$ if one knows the expansion coefficients through first-order.

2. A. L. Stewart, Proc. Phys. Soc. 83, 1033 (1964).

3. See A. Hibbert, Proc. Phys. Soc. 91, 33 (1967) and references there. Also A. Hibbert, Proc. PHys. Soc. 1B, 1048 (1968) for an analogous situation for the Lithium atom.

4. For a discussion of the possibilities in the analogous Rayleigh-Schroedinger perturbation theory see D. F.-T. Tuan, J. Chem. Phys. 46, 2435 (1967).

SEC. XXVIII

1. See for example, R. Grimsrud, J. Chem. Phys. 43 S 59 (1965); A. Dalgarno and G. W. F. Drake, Chem. Phys. Lett. 3, 349 (1969); also E. Brändas and O. Goscinski, Phys. Rev. A 1, 552 (1970). Also the references in [11] sec XI

2. This is a generalization of the procedure described by W. R. Salzman, J. Chem. Phys. 49, 4864 (1968).

3. See a continuing series of papers by J. Midtdal, and co-workers, one of the latest being K. Aashamar, G. Lyslo, and M. Mitdal; J. Chem. Phys. 52, 3324 (1970).

4. D. M. Bishop and A. Macias, J. Chem. Phys. 51, 4947 (1969).

SEC. XXIX

1. A. Dalgarno and A. L. Stewart, Proc. Roy. Soc. A 257, 534 (1960).

2. For a review see A. D. McLachlan, and M. A. Ball, Rev. Mod. Phys. 36, 844 (1964) esp. Section II.

3. For another choice of $\hat{\varphi}_i^{(0)}$ see T. A. Weber, J. H. Weare, and R. G. Parr, J. Chem. Phys. 54, 1865 (1971).

4. M. Cohen, Proc. Phys. Soc. 82, 778 (1963) and references there. He also solves equations arising in the analogous treatment of

restricted Hartree-Fock. C. S. Sharma, J. Chem. Phys. B 2, 1010 (1969) and references there solves analogous equations for various Hartree-Fock schemes; see also R.C. Wilson J Phys B 4, 228; 311 (1971).

5. J. Linderberg, Phys. Rev. 121, 816 (1961).

6. The formalism of AUHF perturbation theory has been extensively discussed and applied in the literature. Two recent papers containing extensive references to the earlier literature see J. Gerratt and I. M. Mills, J. Chem. Phys. 49, 1719 (1968); and R. Moccia, Chem. Phys. Lett. 5, 260 (1970). For a discussion in terms of density matrices see G. Diercksen and R. McWeeny, J. Chem. Phys. 44, 3554 (1966).

7 For a discussion in terms of density matrices see R. McWeeny and G. Diercksen, J. Chem. Phys. 49, 4852 (1969).

8. M. Cohen and A. Dalgarno, Proc. Roy. Soc. A275, 492 (1963).

SEC. XXX

1. See for example V. McKoy and N. W. Winter, Phys. Rev. 48, 5514 (1968).

2. E. Hylleraas, Z. Phys. 65, 209 (1930).

SEC. XXXI

1. See for example R. Goldstein, J. Math. Phys. 8, 473 (1967), also O. Goscinski, Int. J. Q. Chem. 2, 761 (1968) and references there to the earlier considerations of P.-O. Löwdin.

2. T. K. Rebane and P. A. Braun, Int. J. Q. Chem. 3, 1061 (1969).

3. R. A. Buckingham, Proc. Roy. Soc. A160, 94 (1937).

4. J. H. Epstein and S. T. Epstein, J. Chem. Phys. 42, 3630 (1965); T. J. Dougherty, T. Vladimiroff, and S. T. Epstein, J. Chem. Phys. 45, 1803 (1966). See also the recent calculations by T. R. Singh, and W. J. Meath, J. Chem. Phys. 54, 1137 (1971).

SEC. XXXII

1. See P.-O. Löwdin, Int. J. Q. Chem. 4, 231 (1971) and references there.

SEC. XXXIII

1. For a related discussion of higher orders see K. M. Sando and J. O. Hirschfelder, Proc. Nat. Acad. Sci. 52, 434 (1964); (erratum) 67, 356 (1970).

2. W. H. de Jeu, Mol. Phys. 20, 573 (1971).

3. R. N. Mandan, R. W. Haymaker, and R. Blankenbecler, Phys. Rev. 172, 1788 (1968) Sec. II, also reference 1 above.

4. A. Velenik, T. Zivkovic, W, H. de Jeu, and J. N. Murrell, Mol. Phys. 18, 693 (1970) indicate that they are doing calculations along these lines. See also J. D. Power and R. M. Pitzer, Chem. Phys. Lett. 8, 615 (1971) and references there.

5. P. A. Braun and T. K. Rebane, Opt. and Spec. 22, 275 (1967).

SEC. XXXIV

1. O. Sinanoğlu, Phys. Rev. 122, 491 (1961).

2. F. C. Sanders and C. W. Scherr, Phys. Rev. 181, 84 (1969), Appendix.

3. W.H. Miller J. Chem. Phys. 44, 2198 (1966). See also C. S. Sharma, Proc. Phys. Soc. 41, 50 (1967).

SEC. XXXVI

1. L. M. Delves, Proc. Phys. Soc. 92, 55 (1967). See also S. Aranoff and J. Percus, Phys. Rev. 166, 1255 (1968). Similar ideas have been put forward by W. J. Deal, Jr., R. H. Young, and N. R. Kestner, J. Chem. Phys. 49, 3395 (1968).

2. L. J. Shushek and J. B. Knager, Phys. Rev. A3, 1253 (1971).

3. For calculations using (19) see references in footnote B, Sec. XXXI and footnote C, this section.

4. This general variation method was proposed by Delves in an apparantly unpublished manuscript referred to as "D" in Delves ref. 1, where it was also applied using linear trial functions. It was proposed independently starting from the H(00) point of view (see below in text) by Aranoff and Percus, ref. 1.

5. L. M. Delves, Nuc. Phys. 41, 497 (1963).

6. B. Kirtman (Preprint 1971).

APPENDIX C

1. S. T. Epstein, J. Chem. Phys. 49, 1411 (1968); H. S. Kolker, J. Chem. Phys. 49, 3106 (1968); W. H. Miller, J. Chem. Phys. 50, 2758 (1969).

2. L. Rosenberg, L. Spruch, and T. F. O'Malley, Phys. Rev. 118, 184 (1960).

FOOTNOTES

SEC. XXIV

A. Note however that these approximate polarizabilities, etc. will in general have no special variational relationship to their exact counterparts. Thus we know that in general

$$\hat{E} = E + O(\psi - \Psi)^2$$

from which it follows that $\hat{E}$ will be a variational approximation to E also for $\nu = 0$. However unless $\psi - \Psi$ vanishes to some order as $\nu \to 0$ we can draw no conclusions about the relationship between the derivatives of $\hat{E}$ at $\nu = 0$ and those of E

B. Of course if one's sole objective is numerical accuracy it may well be that instead of making a direct expansion, which clearly is the variational analogue of Rayleigh-Schroedinger perturbation theory, one would do better to use some other perturbation technique, or to supplement the expansion by the use of techniques which often speed convergence like Pade approximants or partial fractions. See for example, T. L. Barr, Int. J. Q. Chem. 54, 239 (1971); E. J. Brändas and R. J. Bartlett, Chem. Phys. Lett. 8, 153 (1971); and E. Brändas and O. Goscinski, Phys. Rev. A1, 552 (1970).

SEC. XXV

A. See G. W. F. Drake, and M. Cohen, J. Chem. Phys. 48, 1168 (1968) and references there for a discussion within Rayleigh-Schroedinger perturbation theory $(\hat{E} = E)$. They are concerned among other things with calculating the dipole polarizability of an atom using $1/\omega^2$ expansion techniques. To make contact with our discussion

which assumes that $H^{(0)}$ is independent of ν_1 and ν_2 it is then important to use scaled coordinates and to divide all energies by ω_L^2 so that our H is the actual Hamiltonian divided by ω_L^2.

Then we can identify ν_2 with $1/\omega_L$ and ν_1 with $\mathcal{E}/\omega_L^3$ where $\mathcal{E}$ is the electron field strength. Then the point is that $(\partial^2 E/\partial \nu_1^2)_{\nu_1=0}$ may, in the case of degeneracy for $\nu_2 = 0$, start off with a leading term proportional to $1/\nu_2$. However still, as shown in the reference cited, it may be possible to simplify the calculation of the first correction to this, which now will be a term of zero order in ν_2, by use of interchange theorems. [To be more specific, the second derivative of the actual energy with respect to $\mathcal{E}$ at $\mathcal{E}=0$ is given by $\frac{1}{\omega_L^4}\left(\frac{\partial^2 E}{\partial \nu_1^2}\right)_{\nu_1=0}$, and therefore, in the non-degenerate case starts off like $1/\omega_L^4$. However with degeneracy it may start like $1/\omega_L^3$].

SEC. XXVII

A. With $n=1$ we also of course get a formula for $\hat{E}^{(1)}$ and clearly it will be (XXV-17) instead the simpler (XXV-22). Similarly with higher n values, one gets the simplest formulae only for $\hat{E}^{(2n)}$ and $\hat{E}^{(2n+1)}$.

SEC. XXVIII

A. The calculations of C. W. Scherr and R. E. Knight, Rev. Mod. Phys. 35, 436 (1963) do not fall precisely into this category. See footnote D, Sec. XXXI for more details.

SEC. XXIX

A. UHF in the presence of external fields is often called the "coupled Hartree-Fock approximation." If one neglects the quantity $\hat{f}^{(1)}$ in Eq. (17) below one has the formally simplest version of the "uncoupled Hartree-Fock approximation". See A. Dalgarno, Adv. in Phys. 11, 281 (1962); also P. W. Langhoff, M. Karplus, and R. P. Hurst, J. Chem. Phys. 44, 505 (1966); and footnote C Sec XXXVI Below.

B. Eq. (22) follows automatically only if $\epsilon_k^{(0)} \neq \epsilon_L^{(0)}$. If there is degeneracy of $\epsilon_k^{(0)}$ and $\epsilon_L^{(0)}$ then $(\hat{\varphi}_k^{(0)}, \hat{\varphi}_L^{(1)})$ is left undetermined by (20). If the degeneracy is broken then $(\hat{\varphi}_k^{(0)}, \hat{\varphi}_L^{(1)})$ will be determined by the appropriate higher order equation (the second order equation if the degeneracy is broken in first-order) with a result consistent with (22). If the degeneracy is never broken then (22) may be assumed without loss of generality.

C. If symmetry is a consideration then more precisely "same" means that if the $\hat{\varphi}_L^{(0)}$ and the $\hat{\varphi}_j^{(1)}$ have any symmetries in common then the same basis functions of that symmetry are involved.

SEC. XXX

A. If one keeps on with $\Psi^{(0)} + \nu \Psi^{(1)} + \nu^2 \Psi^{(2)} + \cdots$ then one will of course get to the same place eventually. However one will now find that the equation for $\Psi^{(2)}$ arises by varying $\hat{\Psi}^{(1)}$ and that to derive $T_H^{(4)}$ (defined) below, one will have to go through the same sort of rearranging that we did at the outset of Sec. XXVI.

SEC. XXXI

A. From our derivation it is clear that for the ground state $\hat{J}^{(2)}_D \geq E^{(2)}$. However it is also interesting to give a direct proof. Writing $\hat{\Phi}^{(1)} = a(\psi^{(0)} + \delta)$ where a is some constant one finds, after a bit of rearranging that

$$\hat{J}^{(2)}_D - E^{(2)} = |a|^2 \frac{(\delta, (H^{(0)} - E^{(0)})\delta)(\psi^{(1)}, (H^{(0)} - E^{(0)})\psi^{(1)}) - |(\delta, (H^{(0)} - E^{(0)})\psi^{(1)})|^2}{(\hat{\Phi}^{(1)}, (H^{(0)} - E^{(0)})\hat{\Phi}^{(1)})}$$

If we now denote $\sqrt{H^{(0)} - E^{(0)}}\,\psi^{(1)}$ by ζ and $\sqrt{H^{(0)} - E^{(0)}}\,\delta$ by Γ then since the square root is Hermitian if $E^{(0)}$ is the ground state, the numerator in the fraction can be written

$$(\Gamma, \Gamma)(\zeta, \zeta) - |(\Gamma, \zeta)|^2$$

which is positive by the Schwartz inequality. Since the denominator is obviously positive for the ground state, this completes the proof.

B. An example of this procedure is the uncoupled Hartree-Fock approximation mentioned in footnote A, Sec. XXIX. See also M. Karplus and H. J. Kolker, J. Chem. Phys. 35, 2234 (1961); 38, 1263 (1963). Also the "f-Method" often used in Helium calculations in effect replaces $H^{(0)}$ by the approximate "Steinheimer Hamiltonian" (see for example footnote 5 in A. Dalgarno and S. T. Epstein, J. Chem. Phys. 50, 2837 (1969), also footnote B, Sec. XXXVI below).

C. The result in this form does not appear in the reference 4. However it follows easily from the discussion of Sec. XXXVI below. For some properties of this operator see P.-O. Löwdin, J. Chem. Phys. 43, S 175 (1965) Appendix A.

D. The calculations of Scherr and Knight for $\psi^{(1)}$ and $\psi^{(2)}$ referred to in footnote A, Sec. XXVIII fall into the pattern which we have been

discussing. Namely they use $\hat{\varphi}^{(n)}$ from $\hat{\mathcal{J}}_H^{(n)}$ in $\hat{\mathcal{J}}_H^{(n)}$ instead of $\varphi^{(n)}$ and then in $\hat{\mathcal{J}}_H^{(n)}$ they use $\hat{\varphi}^{(n)}$ and the newly derived $\hat{\varphi}^{(n)}$ instead of $\varphi^{(n)}$ and $\varphi^{(n)}$. If they would use only linear trial functions, with the same basis set in each order then, as we will discuss in the next section, they would be doing perturbation theory within the linear variation method. However they also introduce a non linear parameter which is separately optimized in each order.

SEC. XXXII

A. It is interesting to derive (10) directly from the Dirichlet principle. Thus writing

$$(\hat{\phi}^{(n)}, (H^{(0)} - E^{(0)})\hat{\phi}^{(n)})\hat{\mathcal{J}}_D^{(n)} + (\hat{\phi}^{(n)}, (H^{(n)} - E^{(n)})\varphi^{(0)})(\varphi^{(0)}, (H^{(n)} - E^{(n)})\hat{\phi}^{(n)}) = 0$$

and

$$\hat{\phi}^{(1)} = \sum_k \hat{C}_k \phi_k$$

then $\delta \hat{\mathcal{J}}_D^{(n)} = 0$ all $\delta \hat{C}_k$ yields

$$\hat{\mathcal{J}}_D^{(n)} \sum_L \eta_{kL} \hat{C}_L + (\phi_k, (H^{(n)} - E^{(n)})\varphi^{(0)})(\varphi^{(0)}, (H^{(n)} - E^{(n)})\hat{\phi}^{(n)}) = 0$$

and therefore

$$\hat{C}_L = - \frac{(\varphi^{(0)}, (H^{(n)} - E^{(n)})\hat{\phi}^{(n)})}{\hat{\mathcal{J}}_D^{(n)}} \sum_k (\eta^{-1})_{Lk} (\phi_k, (H^{(n)} - E^{(n)})\varphi^{(0)})$$

or

$$\hat{\phi}^{(n)} = - \frac{(\varphi^{(0)}, (H^{(n)} - E^{(n)})\hat{\phi}^{(n)})}{\hat{\mathcal{J}}_D^{(n)}} \sum_L \sum_k \phi_L (\eta^{-1})_{Lk} (\phi_k, (H^{(n)} - E^{(n)})\varphi^{(0)})$$

Taking the scalar product with $(H^{(n)} - E^{(n)})\varphi^{(0)}$ and cancelling the factor of $(\varphi^{(0)}, (H^{(n)} - E^{(n)})\hat{\phi}^{(n)})$ then yields (10) again. Further we have now determined $\hat{\phi}^{(n)}$ up to a multiplicative constant which is clearly the best we can do. Then using (XXXI-21) and (XXXI-19)

we eventually find (9) again.

B. With one basis function (19) can be derived trivially from the Schwartz inequality (recall also footnote A, Sec. XXXI). Namely (19) then becomes

$$\Omega^{-1} \geq |\phi\rangle (\phi, \Omega\phi)^{-1} \langle\phi|$$

or

$$(g, \Omega^{-1} g)(\phi, \Omega\phi) \geq |(g,\phi)|^2$$

If now we write $g = \sqrt{\Omega^{-1}}\, g$ and $\theta = \sqrt{\Omega}\,\phi$ then this becomes

$$(g, g)(\theta, \theta) \geq |(g, \theta)|^2$$

which is the Schwartz inequality.

SEC. XXXVI

A. In some cases one may also want to model the interactions. For example, in the case of magnetic interactions, and depending on the choice of $H^{(00)}$, this may be necessary if one wishes to have gauge invariance. See for example S. T. Epstein, J. Chem. Phys. 42, 2897 (1965), and A. F. Starace Phys. Rev. A3, 1242 (1971).

B. Suppose that $\chi^{(0)}$ is a UHF wave function calculated from $H^{(0)}$, that $H_1^{(1)}$ is a one-electron operator and that we choose

$$H^{(00)} = F^{(0)}$$

where $F^{(0)}$ is the F of Sec. XI but with $\nu_i = 0$. Then clearly $\phi^{(1)}$ will contain only zero and one-electron excitations of $\chi^{(0)}$. From the definition of $\langle H_1^{(1)} \rangle_0$ the former contribute nothing to

$$(\chi^{(0)}, (H^{(0)} - \langle H_1^{(1)}\rangle_0)\, \phi^{(1)})$$

and, according to Brillouins theorem (Sec. XI) the latter also yield

zero. Hence, using (8), we can rederive the result of Sec. XI that UHF expectation value of one-electron operators are accurate "through first-order".

Turning to (11), $\Phi^{(2)}$ contains zero, one and two electron excitations of $\chi^{(0)}$. However, by the same argument, only the two electron excitations contribute to (11). Since these two electron excitations involve only pairs of first-order spin orbitals (the same ones involved in $\Phi^{(1)}$), it follows that we can compute a first correction to the uncoupled Hartree-Fock polarizability (for that is that $(\partial^2 E/\partial v_i^2)_{\substack{v_i=0 \\ v_{i\neq}=0}}$ is under our assumptions - see footnote A, Sec. XXIX, also footnote C below) just by knowing $\Phi^{(1)}$. See D. F.-T. Tuan, Chem. Phys. Lett. 7, 115 (1970) for a recent summary of calculations based on this result and others like it, and references to the earlier literature.

Also in this paper, the Schuman-Musher (J. Chem. Phys. 49, 4845 (1969) "Geometric Approximation" (first Pade Approximation) is used in a successful attempt to improve on the numerical accuracy of (11).

C. In Footnote A, Sec. XXIX we said that the first order spin orbitals in uncoupled Hartree-Fock satisfy, in the notation of Sec. XXIX,

$$(\hat{f}^{(0)} - \epsilon_i^{(0)})\chi_i^{(1)} + (h^{(1)} - \epsilon_i^{(1)})\hat{\varphi}_i^{(0)} = 0 \qquad (\alpha)$$

However other formally different but physically equivalent equations are sometimes used in the literature. The reason for the ambiguity is as follows: With $H^{(00)} = F^{(0)}$ and $\chi^{(0)} = \hat{\Psi}^{(0)}$, and assuming that $\Phi^{(1)} = \sum_i |\hat{\varphi}_1^{(0)} \cdots \chi_i^{(1)} \cdots \hat{\varphi}_N^{(0)}|$, one finds that Eq. (XXXVI-8) becomes

$$\sum_i |\hat{\varphi}_i^{(0)} \cdots \chi_i \cdots \hat{\varphi}_i^{(N)}| = 0 \qquad (\beta)$$

where χ_i is the left hand side of (α). The general solution of (β) is

$$\chi_i = \sum_j A_{ij} \hat{\varphi}_j^{(0)} \qquad (\gamma)$$

where the A_{ij} can be arbitrary numbers so long as

$$\sum_j A_{jj} = 0 \qquad (\delta)$$

We have made the simplest formal choice, namely $A_{ij} \equiv 0$, however other authors have sometimes made other choices, in order to impose additional constraints on the $\chi_i^{(1)}$, for example the requirement that

$$(\hat{\varphi}_j^{(0)}, \chi_i^{(1)}) = 0 \qquad \text{all } i \text{ and } j$$

Evidently the solution of (γ) differs from that of (α) by the additional terms

$$\sum_{j \neq i} A_{ij} \hat{\varphi}_j^{(0)} / (\epsilon_j^{(0)} - \epsilon_i^{(0)})$$

however clearly the $\phi^{(1)}$'s are identical. Also from (α) $\epsilon_i^{(1)} = (\hat{\varphi}_i^{(0)}, h^{(1)} \hat{\varphi}_i^{(0)})$ while from (β) $\epsilon_i^{(1)} = (\hat{\varphi}_i^{(0)}, h^{(1)} \hat{\varphi}_i^{(0)}) - A_{ii}$ However, from δ, $E^{(1)} = \sum \epsilon_i^{(1)}$ is the same in either case. (We have ignored possible degeneracies. We leave it to the reader to fill in the details in the more general case).

D. This sort of situation is rather common. See for example, J. C. Y. Chen and A. Dalgarno, Proc. Phys. Soc. 85, 399 (1965) in which, using a perturbation approach, they apparantly arrange things so that one doesn't need to know $H^{(00)}$. In fact however this is because $H^{(00)}$ has been fixed implicitly to be the so-called Stetnheimer Hamiltonian

(for simple situations $(H^{(00)} - E^{(00)}) = T - (T\chi^{(0)})/\chi^{(0)}$ where T is the kinetic energy operator). Similarly all applications of which I am aware of the Schwartz method (C. Schwartz, Ann. Phys. (N.Y.) 2, 170 (1959) which also apparantly does not require an $H^{(00)}$, have in fact also been applications of the perturbation method with this same choice of $H^{(00)}$.

E. W. A. Sanders and J. O. Hirschfelder, J. Chem. Phys. 42, 2904 (1965) had earlier observed that (36) and (37) were often equivalent. However they did not relate (36) to the general variation principle (30).